B64

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 December 2000 (21.12.2000)

PCT

(10) International Publication Number
WO 00/77024 A1

(51) International Patent Classification[7]: C07H 21/02, 21/04, A01K 67/00

(21) International Application Number: PCT/US00/16412

(22) International Filing Date: 14 June 2000 (14.06.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/333,229 14 June 1999 (14.06.1999) US

(71) Applicant: NEW YORK BLOOD CENTER, INC. [US/US]; 310 East 67th Street, New York, NY 10021 (US).

(72) Inventors: BELYAVSKY, Alexander V.; 504 East 63rd Street, #10L, New York, NY 10021 (US). POPSUEVA, Anna, E.; Antti Korpintie 4A8, FIN-00600 Helsinki (FI). LUCHINSKAYA, Natalia N.; 56-2 Vavilova Street, #212, Moscow, 117296 (RU).

(74) Agents: GEORGE, Kenneth, P. et al.; Amster, Rothstein & Ebenstein, 90 Park Avenue, New York, NY 10016 (US).

(81) Designated States *(national)*: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *With international search report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 00/77024 A1

(54) Title: CAMELLO GENE FAMILY AND USES THEREOF

(57) Abstract: The present invention provides a purified and isolated nucleic acid encoding a camello protein. The present invention also provides a vector comprising nucleic acid encoding camello, a host cell transformed with the vector and a method for producing recombinant camello protein. In addition, the present invention also provides a purified camello protein. Also provided by the present invention is nucleic acid probes and mixtures thereof specific for camello nucleic acid and antibodies immunoreactive with camello. The present invention also provides a method for screening for agents which bind to the camello protein and the nucleic acid encoding the camello. Finally, the present invention provides a non-human, transgenic model for camello expression.

CAMELLO GENE FAMILY AND USES THEREOF

Background of the Invention

Vertebrate gastrulation involves complex coordinated regulated movements of cells and cell layers to establish the axial structures and the general body plan. Adhesion molecules and the components of extracellular matrix participate in this process. However, other components and detailed mechanisms of the control of gastrulation movements remain largely unknown. For instance, perturbation of cell adhesion by interference with function of different cadherins or extracellular matrix proteins (Kim, *et al.*, *Development* 125, 4681-4691 (1998); Kuhl, *et al.*, *Mechanisms of Development* 54, 71-82 (1996)) has been shown to lead to certain defects in gastrulation. As such, the elucidation of a protein and its nucleic acid involved in cell adhesion may be useful as diagnostic indicators for certain birth defects.

Adhesion molecules mediate cell to cell and cell to matrix interactions and are essential for numerous physiological and pathological processes. The first step of metastasis is the detachment of the tumor cells from the primary tumor and subsequent access to the circulation such as lymph or blood. Although the exact mechanism is unclear at this time, it has been demonstrated that the loss of certain adhesion molecules, such as certain of the cadherins, is associated with the penetration of tumor cells into other tissues and the increased incidence of metastasis, perhaps by facilitating the detachment of the tumor cells from the primary tumor. Accordingly, the elucidation of a protein and its nucleic acid involved in cell adhesion may be useful as a target for anti-metastatic agents.

Summary of the Invention

The present invention is based upon the discovery of a novel gene family, hereinafter denoted "the camello gene family" that the inventors believe is involved in embryogenesis and cell adhesion. This discovery may provide useful targets for anti-metastatic agents, as well as diagnostic indicators for birth defects.

Accordingly, the present invention provides a purified and isolated nucleic acid encoding a camello protein. The present invention also provides a vector comprising this nucleic acid and a host cell transformed by this vector. Also provided by the present invention is a nucleic acid probe which hybridizes to nucleic acid encoding camello, a mixture of nucleic acid probes each of which hybridizes to nucleic acid encoding camello and a kit comprising one or more nucleic acid probes which hybridize to nucleic acid encoding camello.

The present invention also provides a method for producing recombinant camello comprising growing a host cell transformed with a vector comprising nucleic acid encoding camello in culture and recovering the recombinant camello from the culture. The present invention further provides a purified camello protein or an analogue thereof, as well as an agent that binds to the camello protein or its analogue, including but not limited to an antibody immunoreactive with camello or an analogue thereof. In addition, the present invention provides a kit comprising an agent that binds to the camello protein.

The present invention also provides a method for screening an agent that binds to the nucleic acid encoding a camello protein comprising contacting the nucleic acid with an agent of interest and assessing the ability of the agent to bind to the nucleic acid. The present invention further provides for a method for screening an agent that inhibits the expression of the nucleic acid encoding a camello protein comprising contacting a cell transformed with a vector comprising the nucleic acid, and assessing the effect of the agent on expression of the nucleic acid. The present invention still further provides a method for screening for an agent that binds to a camello protein or an analogue thereof comprising contacting the protein with an agent of interest and assessing the ability of the agent to bind to the protein.

In addition, the present invention provides a method for determining the aggressiveness of a tumor in a subject comprising detecting abnormal levels of a camello protein in the tumor relative to normal physiological levels of camello in normal tissue. Further, the present invention provides a method for the diagnosis of birth defects comprising detecting abnormal levels of a camello protein in

embryological tissue relative to normal physiological levels of camello.

The present invention also provides a recombinant viral vector capable of introducing nucleic acid encoding camello into a target cell such that the target cell expresses camello, the vector comprising (a) nucleic acid of or corresponding to at least a portion of the genome of a virus, the portion being capable of infecting the target cell, and (b) nucleic acid encoding a camello protein operably linked to the viral nucleic acid. Finally, the present invention provides a non-human, transgenic animal model comprising mutated nucleic acid encoding camello incorporated into at least some of the somatic cells of the animal.

Additional objects of the present invention will be apparent from the description which follows.

Brief Description of the Figures

Figure 1A depicts the nucleotide sequence of Xenopus camello and Figure 1B depicts the predicted amino acid sequence encoded by the nucleotide sequence of Figure 1A.

Figure 2A depicts the nucleotide sequence of Human camello 1 and Figure 2B depicts the predicted amino acid sequence encoded by the nucleotide sequence of Figure 2A.

Figure 3A depicts the nucleotide sequence of Human camello 2 and Figure 3B depicts the predicted amino acid sequence encoded by the nucleotide sequence of Figure 3A.

Figure 4A depicts a partial nucleotide sequence of Human camello 3 and Figure 4B depicts the predicted amino acid sequence encoded by the nucleotide sequence of Figure 4A.

Figure 5A depicts the nucleotide sequence of Mouse camello 1 and Figure 5B depicts the predicted amino acid sequence encoded by the nucleotide sequence of Figure 5A.

Figure 6A depicts the nucleotide sequence of Mouse camello 2 and Figure 6B depicts the predicted amino acid sequence encoded by the nucleotide sequence of Figure 6A.

Figure 7A depicts the nucleotide sequence of Mouse camello 3 and Figure 7B depicts the predicted amino acid sequence encoded by the nucleotide sequence of Figure 7A.

Figure 8A depicts the nucleotide sequence of Mouse camello 4 and Figure 8B depicts the predicted amino acid sequence encoded by the nucleotide sequence of Figure 8A.

Figure 9A depicts the partial nucleotide sequence of Mouse camello 5 and Figure 9B depicts the predicted amino acid sequence encoded by the nucleotide sequence of Figure 9A.

Figure 10A depicts the nucleotide sequence of Rat camello 1 and Figure 10B depicts the predicted amino acid sequence encoded by the nucleotide sequence of Figure 10A.

Figure 11A depicts the nucleotide sequence of Rat camello 2 and Figure 11B depicts the predicted amino acid sequence encoded by the nucleotide sequence of Figure 11A.

Figure 12A depicts the partial nucleotide sequence of Rat camello 3 and Figure 12B depicts the predicted amino acid sequence encoded by the nucleotide sequence of Figure 12A.

Figure 13A depicts the nucleotide sequence of Rat camello 4 and Figure 13B depicts the predicted amino acid sequence encoded by the nucleotide sequence of Figure 13A.

Figure 14 depicts the alignment of amino acid sequences of camello protein family members. There is a good match between camello consensus sequence and the characteristic motifs of N-acetyltransferase superfamily, positions of which are indicated. The position of the hydrophobic domain is also indicated.

Figure 15 depicts the expression of Xcml during Xenopus development. (a) temporal expression of Xcml mRNA studied by Northern blot analysis, developmental stages are indicated on top. Molecular-size marker is shown at the right. (b-h) spatial pattern of Xcml mRNA expression studied by whole mount in situ hybridization; (dl) dorsal lip; (vl) ventral lip. Expression is first detectable

in the periblastoporal region at the onset of gastrulation, stage 10,5 (b); expression is stronger in the marginal zone at stages 11 (c) and 12 (d). This pattern is preserved until the neurula stage 16 (e). Sagittal sections of Xcml stained Xenopus embryos demonstrate expression of Xcml in deep cells of marginal zone at the beginning of gastrulation movements, stage 10,5 (f) and in the region of contact non-involuted and involuted cells at stage 12 (g). Expression is absent in more deep layers of presumptive mesoderm (g, h), in the cells of outer surface and surface of archenteron (arh); (h) dorsal lip with high magnification.

Figure 16 depicts Xcml overexpression blocking gastrulation movements of cells. Injection of Xcml mRNA in 2 dorsal vegetal blastomeres retards gastrulation (a); blastopore of injected embryos is longer in dorso-ventral direction as a result of suppression of latero-medial movements and intercalation of cells on dorsal side of embryo. At neurula stage injected embryos have short axis and unclosed blastopore (b). Sagittal sections (c-f) of embryos from a show decrease of adhesive properties of cells in injected half of embryo (d, f). Involuted cells form multilayer epithelial structure at the dorsal side (e), but lost this capacity after Xcml overexpression (f). At the neurula stage, injected embryo (g, h) has defect structure of neural plate (np), sometes (som), and disrupted gastrocoel (gc, h).

Figure 17 depicts the effects of Xcml injections on expression pattern of early markers and goosecoid-induced formation of second axis. Expression patterns of actin (a), Xbra (b), Xnot (d, e) marks abnormal position of presumptive materials after Xcml injection. Xcml decreases expression of Pax-6 (c) in posterior part of neural tube and in axial complexes in lateral lips of unclosed blastopore. (f) Injection of gsc in two ventral vegetal blastomeres leads to the formation of full second axis with head structures (eyes, cement glands); (g) co-injection of gsc with Xcml leads to the reduction of head structures.

Figure 18 depicts Xcml protein localized in the organelles of the secretory pathway. (a-c) subcellular localization of Xcml-GFP fusion protein in COS-7 cells studied by confocal microscopy. (a) distribution GFP signal in COS-7 cells; (b)

same as a, but cells were additionally stained with BODIPY TR ceramide, Golgi marker; simultaneous detection of GFP (green) and ceramide (red) signals. (c) COS-7 cells transfected with XcmlDF42L80-GFP construct and stained with BODIPY TR ceramide with simultaneous detection of both signals. (d) COS-1 cells transfected with Xcml-GFP stained with Hoechst that marks nucleus. (e) western blot analysis of Xenopus oocytes injected with C- and N-terminal myc-tagged Xcml and myc-tagged Sizzled as a positive control; M, culture medium; V, vesicular fraction; C, cytoplasmic fraction.

Figure 19 depicts an example of the blastomere aggregation assay for analysis of Xcml function. The number of cells in each aggregate class is indicated on the horizontal axis, and the percentage of cells in each aggregate class is indicated on the vertical axis. The results for injection of identical amounts of Xcml and XcmlA31Fr (A3) RNA are compared. P values on the horizontal axis indicate the probability that the difference between Xcml and the negative control is non-significant. The data demonstrate that Xcml injection substantially reduces blastomere aggregation, since the percentage of single cells after Xcml injection increased 4-fold, while the percentage of large (more than 10 cells) aggregates decreased more than 6-fold in this example.

Figure 20 depicts the blastomere aggregation assay for Hcml1 RNA. Graph details are the same as in Figure 19. On the horizontal axis, the following aggregate size classes are shown: 1: single cells; 2: 2-4 cells; 3: 5-7 cells; 4: 8-10 cells; and 5: more than 10 cells. Compared are the effects on Ca-induced blastomere re-aggregation of injection of identical amounts of Xcml, Hcml1 (Hum), or XcmlA31Fr (A3) RNA. The data indicate that Hcml1 (like Xcml) substantially reduces adhesion of blastomeres. After Hcml1 injection, the number of single cells is increased approximately 2.5-fold, whereas the number of cells in large (more than 10 cells) aggregates is reduced more than 5-fold compared to the negative control (A3).

Detailed Description of the Invention

The present invention provides a purified and isolated nucleic acid encoding a camello protein. As used herein, the nucleic acid may be genomic DNA, cDNA, RNA or antisense RNA and includes nucleic acid derived from any species, *e.g.*, human, rat, goat, pig, mouse, frog and cow. Due to the degeneracy of the genetic code, the nucleic acid of the present invention also includes a multitude of nucleic acid substitutions which will encode camello. The nucleic acid from the frog preferably encodes the amino acid sequence for Xenopus camello (Xcml) as shown in Figure 1B, and more preferably comprises the nucleotide sequence as shown in Figure 1A. The nucleic acid from a human preferably encodes the amino acid sequences for human camello shown in Figures 2B (Hcml1), 3B (Hcml2) or 4B (Hcml3), and more preferably comprises the nucleotide sequence shown in Figures 2A, 3A or 4A, respectively. The nucleic acid from the mouse preferably encodes for the amino acid sequences for mouse camello as shown in Figures 5B (Mcml1), 6B (Mcml2), 7B (Mcml3), 8B (Mcml4) or 9B (Mcml5), and more preferably comprises the nucleotide sequence shown in Figures 5A, 6A, 7A, 8A or 9A, respectively. The nucleic acid for the rat preferably encodes for the amino acid sequences for rat camello as shown in Figures 10B (Rcml1), 11B (Rcml2), 12 B (Rcml3) or 13B (Rcml4), and more preferably comprises the nucleotide sequence set forth in Figures 10A, 11A, 12A, or 13A, respectively.

The present invention also includes nucleic acid sequences that are at least 80%, preferably at least 85%, more preferably at least 90%, and most preferably at least 95%, homologous with each of the nucleic acid sequences set forth above. In addition, the present invention provides the nucleic acid encoding the camello protein having one or more mutations resulting in the expression of a non-functional or mutant protein, or in lack of expression altogether. The mutation may be one or more point, insertion, rearrangement or deletion mutations or a combination thereof.

The present invention further provides a vector which comprises nucleic acid encoding a camello protein. Such vectors may be constructed by inserting nucleic acid encoding camello into suitable vector nucleic acid. The term

"inserted" as used herein means the ligation of a foreign DNA fragment and vector DNA by techniques such as the annealing of compatible cohesive ends generated by restriction endonuclease digestion or by use of blunt end ligation techniques. Other methods of ligating DNA molecules will be apparent to one skilled in the art. Vectors may be derived from a number of different sources. They can be plasmids, viral-derived nucleic acids, lytic bacteriophage derived from phage lambda, cosmids or filamentous single-stranded bacteriophages such as M13. Depending upon the type of host cell into which the vector is introduced, vectors may be bacterial or eukaryotic. Bacterial vectors are derived from many sources including the genomes of plasmids and phage. Eukaryotic vectors are also constructed from a number of different sources, *e.g.*, yeast plasmids and viruses. Some vectors, called shuttle vectors, are capable of replicating in both bacteria and eukaryotes. The nucleic acid from which the vector is derived is usually greatly reduced in size so that only those genes essential for its autonomous replication remain. The reduction in size enables the vectors to accommodate large segments of foreign DNA. Examples of suitable vectors into which the nucleic acid encoding the camello protein can be inserted include but are not limited to pBR322, pUC18, pUC19, pHSV-106, pJS97, pJS98, M13mp18, M13mp19, pSPORT 1, pGem, pSPORT 2, pSV●SPORT 1, pBluescript II, λZapII, λgt10, λgt11, λgt22A, and λZIPLOX. Other suitable vectors are obvious to one skilled in the art.

The vector of the present invention may be introduced into a host cell and may exist in integrated or unintegrated form within the host cell. When in unintegrated form, the vector is capable of autonomous replication. The term "host cell" as used herein means the bacterial or eukaryotic cell into which the vector is introduced. As used herein, "introduced" is a general term indicating that one of a variety of means has been used to allow the vector to enter the intracellular environment of the host cell in such a way that it exists in stable and expressable form therein.

Some bacterial and eukaryotic vectors have been engineered so that they are capable of expressing inserted nucleic acids to high levels within the host cell. Such vectors utilize one of a number of powerful promoters to direct the high

level of expression. For example, in vectors for the expression of a gene in a bacterial host cell such as E. coli, the lac operator-promoter or the tac promoter are often used. Eukaryotic vectors use promoter-enhancer sequences of viral genes, especially those of tumor viruses. Expression can be controlled in both bacterial and eukaryotic cells using inducible promoters such as the lac operator-promoter in E. coli or metallothionine or mouse mammary tumor virus promoters in eukaryotic cells. As used herein, "expression" refers to the ability of the vector to transcribe the inserted nucleic acid into mRNA so that synthesis of the protein encoded by the inserted nucleic acid can occur.

Vectors suitable for the expression of the nucleic acid encoding camello in a host cell are well known to one skilled in the art and include pET-3d (Novagen), pProEx-1 (Life Technologies), pFastBac 1 (Life Technologies), pSFV (Life Technologies), pcDNA II (Invitrogen), pSL301 (Invitrogen), pSE280 (Invitrogen), pSE380 (Invitrogen), pSE420 (Invitrogen), pTrcHis A,B,C (Invitrogen), pRSET A,B,C (Invitrogen), pYES2 (Invitrogen), pAC360 (Invitrogen), pVL1392 and pVl1392 (Invitrogen), pCDM8 (Invitrogen), pcDNA I (Invitrogen), pcDNA I(amp) (Invitrogen), pZeoSV (Invitrogen), pcDNA 3 (Invitrogen), pRc/CMV (Invitrogen), pRc/RSV (Invitrogen), pREP4 (Invitrogen), pREP7 (Invitrogen), pREP8 (Invitrogen), pREP9 (Invitrogen), pREP10 (Invitrogen), pCEP4 (Invitrogen), pEBVHis (Invitrogen), and λPop6. Other vectors would be apparent to one skilled in the art.

Vectors may be introduced into host cells by a number of techniques known to those skilled in the art, *e.g.*, electroporation, DEAE dextran, cationic liposome fusion, protoplast fusion, DNA coated-microprojectile bombardment, and infection with recombinant replication-defective retroviruses. The term "transformation" denotes the introduction of a vector into a bacterial or eukaryotic host cell. As such, it encompasses transformation of bacterial cells and transfection, transduction and related methods in eukaryotic cells.

Any one of a number of suitable bacterial or eukaryotic host cells may be transformed with the vector of the present invention. Examples of suitable host cells are known to one skilled in the art and include but are not limited to bacterial cells such as E. coli strains c600, c600hfl, HB101, LE392, Y1090,

JM103, JM109, JM101, JM107, Y1088, Y1089, Y1090, Y1090(ZZ), DM1, PH10B, DH11S, DH125, RR1, TB1 and SURE, Bacillus subtilis, Agrobacterium tumefaciens, Bacillus megaterium; and eukaryotic cells such as Pichia pastoris, Chlamydomonas reinhardtii, Cryptococcus neoformans, Neurospora crassa, Podospora anserina, Saccharomyces cerevisiae, Saccharomyces pombe, Uncinula necator, cultured insect cells, cultured chicken fibroblasts, cultured hamster cells, cultured human cells such as HT1080, MCF7, 143B and cultured mouse cells such as EL4 and NIH3T3 cells.

The present invention also provides a method for producing a recombinant camello protein comprising growing a host cell transformed with a vector encoding camello in culture and recovering recombinant camello. As used herein the term "recombinant" refers to camello produced by purification from a host cell transformed with a vector capable of directing its expression to a high level. A variety of methods of growing host cells transformed with a vector are known to those skilled in the art. The type of host cell, *i.e.*, whether the host cell is bacterial or eukaryote, is the primary determinant of the method to be utilized and the optimization of specific parameters relating to such factors as temperature, trace nutrients, humidity, and growth time. Depending on the vector, the host cells may have to be induced by the addition of a specific compound at a certain point in their growth cycle in order to initiate expression of the nucleic acid of the present invention. Examples of compounds used to induce expression of the nucleic acid of the present invention are known to one skilled in the art and include but are not limited to IPTG, zinc and dexamethasone. Using standard methods of protein isolation and purification, such as ammonium sulfate precipitation followed by dialysis to remove salt, followed by fractionation according to size, charge of the protein at specific pH values, affinity methods, etc., recombinant camello may be extracted from suitable host cells transformed with vector capable of expressing the nucleic acid encoding camello.

The present invention also provides a purified camello protein and analogues thereof and includes camello isolated from tissue obtained from a subject or recombinantly produced as described above. As used herein

"analogues" may be any protein having functional similarity to the camello protein, that also possesses certain regions that are conserved among the Camello family members (*e.g.*, the central hydrophobic domain). Preferably, the camello protein from the frog preferably comprises the amino acid sequence for Xenopus camello (Xcml) as shown in Figure 1B. Preferably, the camello protein from the human comprises the amino acid sequences shown in Figures 2B (Hcml1), 3B (Hcml2) or 4B (Hcml3). The camello protein for the mouse preferably comprises the amino acid sequences as shown in Figures 5B (Mcml1), 6B (Mcml2), 7B (Mcml3), 8B (Mcml4) or 9B (Mcml5. The camello protein for the rat preferably comprises the amino acid sequences shown in Figures 10B (Rcml1), 11B (Rcml2), 12 B (Rcml3) or 13B (Rcml4). The camello protein also includes amino acid sequences that are at least 70%, preferably at least 75%, more preferably at least 80%, and most preferably at least 90% homologous with each of the amino acid sequences set forth above. The present invention also includes a non-functional camello protein, *i.e.*, camello which is inactive or only has minimal effects *in vivo*. The non-functional camello protein may have one or more deletions or substitutions of its amino acid sequence that results in the camello protein losing its functionality.

The present invention also provides for agents that bind to the camello protein and analogues thereof, as well as the non-functional camello protein. The agent may be a antibody, a nucleic acid, a protein, a peptide, DNA, RNA, mRNA, antisense RNA, a drug or a compound. Agents that bind to the camello protein or an analogue thereof may be identified or screened by contacting the protein with the agent of interest and assessing the ability of the agent to bind to the protein. Agents that bind to the camello protein may act to inhibit metastasis by inhibiting the anti-adhesion effects of camello expression and, therefore, may be useful as chemotherapeutic agents for cancer and tumor treatment. Such agents also may be useful for the treatment or prevention of birth defects.

Antibodies immunoreactive with camello or analogues thereof include antibodies immunoreactive with non-functional camello protein. The antibodies of the present invention may be monoclonal or polyclonal and are produced by techniques well known to those skilled in the art, *e.g.*, polyclonal antibody can be

produced by immunizing a rabbit, mouse, or rat with purified camello and monoclonal antibody may be produced by removing the spleen from the immunized rabbit, mouse or rat and fusing the spleen cells with myeloma cells to form a hybridoma which, when grown in culture, will produce a monoclonal antibody. Labeling of the antibodies of the present invention may be accomplished by standard techniques using one of the variety of different chemiluminescent and radioactive labels known in the art. The antibodies of the present invention may also be incorporated into kits which include an appropriate labeling system, buffers and other necessary reagents for use in a variety of detection and diagnostic applications.

The present invention provides for agents that bind to a nucleic acid encoding camello protein. Suitable agents include but are not limited to a nucleic acid, a protein, a peptide, DNA, RNA, mRNA, antisense RNA, a drug or a compound. Preferably, the agents inhibit expression of the camello nucleic acid. Such agents may be discovered by a method for screening for an agent that binds to the nucleic acid of camello comprising contacting the nucleic acid with an agent of interest and assessing the ability of the agent to bind to the nucleic acid. An agent that inhibits the expression of the nucleic acid encoding the camello protein may be screened by contacting a cell transformed with a vector comprising the nucleic acid, and assessing the effect of the agent on expression of the nucleic acid. Agents that bind to the nucleic acid encoding camello may act to inhibit metastasis of tumors by inhibiting the anti-adhesion effects of camello expression.

The present invention also provides nucleic acid probes and mixtures thereof which are hybridizable to the nucleic acid encoding the camello protein. Such probes may be prepared by a variety of techniques known to those skilled in the art such as PCR and restriction enzyme digestion of camello nucleic acid or by automated synthesis of oligonucleotides whose sequences correspond to selected portions of the nucleotide sequence of the camello nucleic acid using commercially available oligonucleotide synthesizers such as the Applied Biosystems Model 392 DNA/RNA synthesizer. The nucleic acid probes of the present invention may also be prepared so that they contain one or more point,

insertion, rearrangement or deletion mutations or a combination thereof to correspond to mutations of the camello gene. The nucleic acid probes of the present invention may be DNA or RNA and may vary in length from about 8 nucleotides to the entire length of the camello nucleic acid. Preferably, the probes are 8 to 30 nucleotides in length. Labeling of the nucleic acid probes may be accomplished using one of a number of methods known in the art, *e.g.*, PCR, nick translation, end labeling, fill-in end labeling, polynucleotide kinase exchange reaction, random priming, or SP6 polymerase (for riboprobe preparation) and one of a variety of labels, *e.g.*, radioactive labels such as ^{35}S, ^{32}P, or ^{3}H or nonradioactive labels such as biotin, fluorescein (FITC), acridine, cholesterol, or carboxy-X-rhodamine (ROX). Combinations of two or more nucleic probes corresponding to different or overlapping regions of the camello nucleic acid may also be included in kits for use in a variety of detection and diagnostic applications.

The present invention also provides a method for diagnosing developmental defects in an embryo or fetus associated with abnormal expression in the subject's cells. Abnormal expression of camello may be associated with defects in gastrulation. Gestational defects in an embryo or fetus resulting from an increased or decreased expression of camello may be diagnosed by nucleic acid hybridization and/or immunological techniques well known in the art. For example, nucleic acid hybridization using mRNA extracted from cells and camello nucleic acid probes can be used to determine the concentration of camello mRNA present in the cell and the concentration thus obtained compared to the value obtained for cells which exhibit a normal level of camello activity. Isolation of RNA from cells is well known in the art and may be accomplished by a number of techniques, *e.g.*, whole cell RNA can be extracted using guanidine thiocyanate; cytoplasmic RNA may be prepared by using phenol extraction methods; and polyadenylated RNA may be selected using oligo-dT cellulose. Alternatively, the concentration of camello in the cell may be determined from binding studies using antibody immunoreactive with camello. Gestational defects resulting from mutations in the nucleic acid encoding camello may be detected by one of a number of methods known in the art, *e.g.*, hybridization analysis of nucleic acid

extracted from a sample of tissue or cells from a subject using nucleic acid probes designed to detect the presence of mutations in the nucleic acid encoding camello. Alternatively, the defect may be detected using antibody immunoreactive with non-functional camello and standard immunological detection techniques such as Western blotting.

Increased expression of camello in cancer or tumor cells, which may be indicative of increased metastasis or aggressiveness of the tumor, may be detected by nucleic acid hybridization and/or immunological techniques well known in the art. For example, nucleic acid hybridization using mRNA extracted from cells and camello nucleic acid probes can be used to determine the concentration of camello mRNA present in the cell and the concentration thus obtained compared to the value obtained for cells which exhibit a normal level of camello activity. Alternatively, the concentration of camello in the cell may be determined from binding studies using antibody immunoreactive with camello.

Finally, the method of the present invention also provides a non-human animal model for the study of camello expression. The animal model of the present invention comprises a non-human, transgenic animal having nucleic acid encoding the camello protein incorporated into at least some of the somatic cells of the animal. The effect of the expression of the camello protein also may be studied by overexpressing or underexpressing the protein using suitable promoters and regulators known in the art. It is also within the confines of the present invention that a nucleic acid sequence having one or more mutations may be introduced into the animal model that result in the expression of a non-functional or mutant protein. Nucleic acid encoding mutated camello may be integrated into the germ line of a non-human animal such as a mouse, rat, goat, sheep, or other species in order to obtain a transgenic animal. Expression of the incorporated nucleic acid may be restricted to certain tissues in the transgenic animal by the utilization of tissue-specific promoters. Methods of making transgenic animals are well known in the art. For example, DNA encoding mutated camello can be inserted into the genome of a replication-defective virus such as HSV, or a retrovirus or transposon, and the resultant construct injected into embryonic stem cells. Transgenic animals may also be made by injecting

DNA encoding mutated camello into the male pronucleus of a fertilized egg of a nonhuman animal, transplanting the "transgenic embryo" into a pseudopregnant female and then analyzing offspring for the presence of the injected DNA in their genome. Other methods of producing transgenic mice would be apparent to one skilled in the art.

The present invention is described in the following Experimental Details Section which is set forth to aid in the understanding of the invention, and should not be construed to limit in any way the invention as defined in the claims which follow thereafter.

Experimental Details Section

1. Materials and Methods

Molecular analysis of *Xcml* and mammalian camello family members

N4 *Xcml* cDNA fragment, isolated using Gene Expression Fingerprinting procedure, was used as a probe for screening *Xenopus laevis* stage 10,5 embryo cDNA library. pBluescript SK(-) plasmids were excised from positive clones using R408 helper phage, and the largest clone 1,2 kb long was sequenced in both directions. EST clones containing murine, rat and human camello family sequences were obtained from Genome Systems, Inc. (St. Louis, MO) and ATCC and sequenced using flanking and gene-specific primers.

Plasmid constructs and site-directed mutagenesis

For microinjection experiments the *Xcml* open reading frame was PCR amplified with Advantage cDNA polymerase mix and inserted into BamH1/Xba1-cleaved pCS2+ vector. Xcml constructs fused – or C-terminally with six tandemly-repeated copies of myc epitope (myc-tag) were produced by in-frame insertion of PCR amplified Xcml open reading frame into Xho1/Xba1- or BamH1-digested pCS2+MT vector. For Xcml constructs C-terminally fused with Green Flourescent Protein PCR fragments containing intact protein coding sequence or sequence with deletion of hydrophobic domains were cloned in-frame into Xho1-BamH1 sites of pEGFP-N1 vector (Clontech).

Constructs of mutated Xcml protein were created using site-directed

mutagenesis by inverse PCR. Xcml-pCS2+ circular plasmid nicked by DNAse I as described was used as a template. Amplifications were carried out using the Advantage cDNA PCR kit (Clontech) for 10 cycles (95°C, 30 seconds; 60°C, 30 seconds; 68°C, 4 minutes). The amplified fragments were gel purified and self-ligated. XcmlA31F mutant contained a frameshift after Ala31 and a translation stop five amino acids further downstream. XcmlA32S and XcmlQ147S had stop-codons after Ala32 and Gln147, respectively. In the XcmlDF42L80 mutant an internal hydrophobic domain between Arg41 and Glu81 was deleted. All constructs and mutants were checked by sequencing.

Northern blot analysis

Isolation of total and poly(A)$^+$ RNA from embryos were performed as described. For Northern analysis poly(A)$^+$ RNA was separated in a 1.2 % formaldehyde-agarose gel and transferred by capillary blotting onto Hybond-N nylon membrane according to manufacturer instructions. Blot was probed with [^{32}P]dATP-labeled *Xcml* and washed in stringent conditions.

In situ hybridization

Whole-mount in situ hybridization was performed according to Harland (1991) using digoxigenin-labeled antisense RNA probes synthesized from *Xcml* plasmids using T7 RNA polymerase.

RNA synthesis and microinjection

Synthetic capped sense mRNAs were produced using the Ambion Message Machine SP6 kit using corresponding linearized plasmids. *Xenopus* embryos were obtained by *in vitro* fertilization, chemically degelled with 2% cysteine hydrochloride (pH 8.0) at the 2-cell stage, washed with 0.1x MMR (1xMMR: 100mM NaCl, 2mM KCl, 1mM MgSO4, 2mM CaCl2, 5mM Hepes, pH 7.6, 0.1 mM EDTA) and transferred to 1/3x MMR supplemented with 4% Ficoll type 400 (Sigma). Capped mRNA in 4.6 nl of RNAse-free water was injected in embryos at the 8-cell stage. At the mid-blastula stage embryos were placed in 0.1x MMR. Staging was performed according to Nieuwkoop and Faber (1975).

Western blot analysis

Manually defolliculated oocytes were injected in OR2 medium (82.5 mM NaCl, 2.5 mM KCl, 1mM CaCl2, 1 mM MgCl2, 1 mM Na2HPO4, 5 mM HEPES, pH 7.6) with 30 ng of mRNA in 28 nl of water and cultured overnight at room temperature in 0.5x MMR, 0.5 mg/ml BSA, 50 units/ml penicillin, 50 mg/ml streptomycin (10 ml per oocyte). Culture medium was collected and acetone-precipitated. Oocytes were fractionated into cytosolic and vesicle fractions. Proteins (20 mg per lane) were separated by SDS-PAGE, transferred to nitrocellulose membrane by electroblotting and probed with rabbit antibodies against myc-tag followed by goat anti-rabbit IgG secondary antibodies conjugated with horseradish peroxidase (Amersham). The protein bands were visualized using enhanced chemiluminescence.

Cell culture

For cell-localization findings COS-1 cells were transfected with 10 mg of Xcml-pEGFP-N1 plasmid or pEGFP-N1 using the calcium phosphate technique (Graham and Van Der Eb, 1973) and cultured up to 3 days in DMEM with 10 % FBS medium. For nuclear staining 0.25 mg/ml of Hoechst was added to the culture medium two hours before examination. Cells were examined under FITC filter on a Leica microscope equipped with photo camera.

Confocal microscopy imaging

Cells growing on glass coverslips were transiently transfected with either Xcml-GFP expression construct or XcmlF2-GFP using FuGene 6 transfection kit (Boehringer Mannheim) 36 h prior to fluorescence analysis. Golgi apparatus was stained by treatment of cells with 0.5 mM BODIPY TR ceramide (Molecular Probes) for 1h. After loading, the cell were washed twice with and kept in Dulbecco/PBS solution containing 20 mM HEPES, pH 7.4 at room temperature for 20 min prior to the experiments. The fluorescence in living cells was analyzed using a Bio-Rad MRC-1024 confocal microscope equipped with an argon-krypton laser.

Expression of the members of camello family

The effect on cell adhesion of overexpression of Xcml and a human member of the camello family (Hcml1) was studied using blastomere aggregation assay. For the aggregation assay, 2 ng of Xcml or XcmlA31Fr (mutant with the frameshift after Ala31, which served as a negative control) mRNA were injected into the animal pole of blastomeres at the 4-cell stage. Animal caps were isolated at stage 8, and the blastomeres were dissociated in calcium/magnesium-free MMR medium (100 mM NaCl; 2 mM KC1; 5 mM Hepes, pH 7.6; 0.1 mM EDTA) by passing several times through the plastic tip. Calcium was added to the medium, to a concentration of 2 mM, and blastomeres were allowed to aggregate on a horizontal rotary shaker at 60 rpm in 35-mm dishes coated with 1% agarose (10 caps per dish). After incubation for 30 to 40 min, cells were fixed by addition of formaldehyde to 4%.

The aggregates were divided into five size classes and quantified. The size classes consisted of: a) single cells; b) 2-4 cells; c) 5-7 cells; d) 8-10 cells; and e) more then 10 cells per aggregate. Differences in the total number of cells in aggregates of each size class after injection of Xcml and XcmlA31Fr were evaluated in eight experiments. The Wilcoxon test was used for statistical comparisons. P values less than 0.05 were accepted as indicating statistically-significant differences between the two samples. For analysis of Hcml1 influence on cell adhesion, effects of injection of 2 ng of Hcml1 RNA or of XcmlA31Fr RNA were compared.

2. Results and Discussion

To identify genes potentially involved in regulation of gastrulation, Gene Expression Fingerprinting technique (Ivanova and Belyavsky, *Nucl. Acid Res.* 23: 2954-2958 (1995)) was used to search for genes expressed differentially in subregions of Xenopus gastrula embryos. One of the identified sequences (N4) was found to be expressed specifically in the dorsal and ventral marginal zones (Ivanova, *et al.*, *Dokl. Acad. Nauk* 359:116-119 (1998)) at the beginning of gastrulation, and its detailed study is described herein. A cDNA clone isolated from gastrula library encodes the predicted protein 219 amino acids long (Fig.

1B) containing an internal 40-amino acid long hydrophobic region with a short hydrophilic stretch in the middle suggesting that the protein can be membrane associated. At the same time, no N-terminal hydrophobic leader peptide sequence typical for transmembrane proteins could be found. Due to the characteristic hydrophobicity profile of the encoded protein the gene was named *camello* (Spanish for camel).

Searches in the EST database revealed four murine (Mcml1-4), two rat (Rcml1,2) and one human (Hcml1) non-identical cDNA sequences encoding putative proteins with significant homology to Xenopus camello (Xcml) and to each other. A second human putative member of this family was identified in the Huntington gene region whereas TSC501 gene (Ozaki, *et al.*, *J. Hum. Genet.* 43, 255-258 (1998)) is virtually identical to the human Hcml1 gene. Deduced amino acid sequences of the mammalian camello family are shown on Fig. 14. At amino acid level, Xcml is 37% identical to human/mouse, whereas the human-mouse identity is 60% with conservative replacements. Mammalian homologues also demonstrate a striking similarity to Xcml at the structural level, including the presence of hydrophobic domain, its length, organization and the distance from the N-terminus. Moreover, C-terminal regions of Xenopus and mammalian members of camello family demonstrate statistically significant homology to the different members of the large family of N-acetyltransferases present in bacteria, fungi and animals (Lee, *et al.*, *J. Biol. Chem.* 263:14948-14955 (1988); Hintermann, *et al.*, *FEBS Lett.* 375:148-150 (1995); Ebisawa *et al.*, *Eur. J. Biochem.* 228:129-137 (1995)). The maximum degree of identity of camello family members to N-acetyltransferases is 25-30%, fairly similar to the homology between different N-acetyltransferase groups (Coon, *et al.*, *Science* 270:1681-1683 (1995)). Two structural domains responsible for Ac CoA binding (domain A) and acetyl group transfer (domain B) were identified in N-acetyltransferases (Schulz, *Curr. Opin. Struct. Biol.* 2:61 (1992)). All *camello* family members match well the consensus motifs in both A and B N-acetyltransferase domains. It should be noted that no N-acetyltransferases with extended hydrophobic regions have been reported, and the only member of this family with a demonstrated role

in embryo development is the Hookless participating in plant morphogenesis. On the basis of protein sequence analysis it is suggested that the camello family is a novel and highly distinct subgroup of N-acetyltransferases.

Temporal pattern of *Xcml* gene expression was studied by the Northern blot analysis (Fig. 15A). *Xcml* gene encodes a c.a. 1.4 kb transcript that appears after MBT, reaches its expression maximum at the stage 10 and continues to be expressed at similar levels until at least stage 27.

Whole-mount in situ hybridization using a digoxygenine-labeled *Xcml* RNA antisense probe (Fig. 15B) revealed that the first weak signal appears in the marginal zone of embryo at the beginning of gastrulation (stage 10), in the region of presumptive chordamesoderm. *Xcml* is expressed in deep cells of this zone. Bottle cells -the leading cells of dorsal lip- are not stained. Larger magnification reveals mosaic staining of marginal zone with many cells not stained. At stages 11,5 and 12 *Xcml* message is expressed in the same ring of deep cells around the closed blastopore. During gastrulation marginal zone cells initially expressing *Xcml* involute, perform convergent-intercalation movements and form axial structures (chorda and somites) at the dorsal side of embryo. However, hybridization data demonstrate that whereas *Xcml* is expressed in the surface cells of periblastopore region, these cells cease to express gene after they involute inside the embryo. This expression pattern is substantially different from that of other genes expressed in presumptive mesoderm, most of which continue to be expressed after involution. At late neurula and tailbud stages, Xcml transcripts are found in the deep mass of cells lying ventrally and laterally to the chordoneural hinge.

To investigate the role that Xcml might play during early development, *in vitro* synthesized *Xcml* mRNA was injected into equatorial region of dorsal or ventral blastomeres of 8-cell stage embryos. With dorsal injection, development proceeded normally until the late blastula, but during gastrulation the involution of mesoderm in the majority (up to 70%) of injected embryos was greatly inhibited. Blastopore closure did not occur completely (Fig. 16, Table 1 below) and until neurula stages most of the embryos keep open blastopores of different

sizes; in some abnormal embryos blastopore closure did not occur at all. In these cases, mesodermal cells during epiboly spread along the big blastopore resulting in two bands of axial tissue on each side of the blastopore. The multilayer accumulation of mesodermal cells in ventro-lateral region of the blastopore was detected on the sagittal sections of dorsally injected embryos. The suppression of radial intercalation movements led to the phenotype with shortened antero-posterior axis with severely truncated head structures and neural plate. Little if any developmental defects were observed in embryos injected with the same amounts of actin mRNA. Introduction of the frame-shift after Ala31 or the stop-codon after Ala32 (constructs XcmlA31F and XcmlA32S, respectively) resulted in complete elimination of developmental abnormalities demonstrating the specificity of effects produced by camello RNA .

When ventral blastomeres were injected, embryos appeared normal until the late gastrula stages. Embryos successfully formed ventral lip, but mesodermal cells accumulated in the lateral region which became apparent in asymmetrically injected embryos with curved posterior parts of axial complexes (Fig. 16).

To study in more detail the developmental defects produced by Xcml overexpression, whole-mount in situ hybridization of injected embryos with mesodermal and neural tissue markers such as Xbra, Xnot, b-tubulin, eng, Pax6, gsc, chr, nog, BMP4 was performed. Observed patterns were fully compatible with morphological changes caused by defects in gastrulation (Fig. 16). Therefore, overexpression of Xcml, apart of mechanistic effects, seems to induce little if any changes in gene expression or in the determination of the cell layers.

Dorsal overexpression of Xcml mutant protein with deletion of N-acetyltransferase domain but intact N-terminal two thirds had essentially no effect on gastrulation indicating that the deleted domain is necessary for the function of the protein. At the same time, overexpression of the Xcml mutant (XcmlDF42L80) devoid of the entire hydrophobic domain showed moderate, two- to three-fold, reduction in the percentage of gastrulation defects compared to the intact protein, suggesting that the hydrophobic domain, although essential, is not indispensable for *camello* function.

Table 1. Xcml overexpression inhibits gastrulation movements and induction of the ectopic axis

	n	Abnormalities of development (%)	Complete secondary axis (%)	Reduced secondary axis (%)
Xcml	125	77		
	160	58		
Actin	159	2		
XcmlA31F	72	0		
XcmlJ	79	0		
XcmlDF42L80	89	29		
Mcml 1	39	50		
Hcml1	74	50		
Goosecoid (60 pg)	25		50	50
Goosecoid (60 pg) + Xcml (1 ng)	37		25	75

For experiments where inhibition of gastrulation movements were examined, 8-cell stage embryos were injected in two dorsal vegetal blastomeres with 2 ng per embryo of the indicated RNAs. In assay of ectopic axis induction, the same stage embryos were injected in two ventral vegetal blastomeres, and secondary axes were scored at the tailbud stage. Duplicated axes were scored as complete when showing cement gland and eyes, and as reduced when lacking both features.

Sections were prepared to study overexpression Xcml on cell morphology. There are large spaces between cells and cavities in injected dorsal part of embryos as compared with ventral part (Fig. 17b) and intact embryos (Fig. 17c). Cells change from polygonal shape to elongate. Involuted cells of intact embryos form multilayer epithelial structure at dorsal side (future chorda and somites) (Fig. 17d), epithelial sheets form archenteron. Overexpression Xcml disorders

epithelial structures (Fig. 17e). Observations allow to suppose about decreasing adhesion ability by descendants of injected Xcml blastomeres. Changes of morphogenetic behavior of cells through gastrulation led to morphology defects at neurula stage: abnormal structure of neural plate, somites, asymmetric position and disruption of integrity of gastrocoel.

Ectopic expression of goosecoid on the ventral side of embryo induces a massive cell movement at the early gastrula stage toward the anterior of the embryo and formation of second axis (Niehrs, *et al.*, *Cell* 72:491-503 (1993)). Xcml evidently antagonizes this action of goosecoid since co-injection of *Xcml* and *goosecoid* mRNAs in two ventral blastomeres led to the decrease of formation of complete secondary axes from 60% in embryos injected with gsc alone to 27% in co-injected embryos. This result presents an additional evidence for an inhibitory effect of Xcml overexpression on gastrulation movements.

The possible function of mammalian members of camello family was studied by injection of RNA of Mcml4 and Hcml1 genes into Xenopus dorsal blastomeres. In both cases the nature and magnitude of developmental effects were quite similar to those observed with control injections of Xcml RNA (Table 1, above) suggesting the similarity of mechanisms of action and possibly *in vivo* functions of mammalian and Xenopus camello proteins.

To determine the intracellular localization of Xcml protein, the inventors performed the confocal microscopy of COS-7 cells transfected with the Xcml-GFP fusion expression construct. The majority of fluorescent signal was found in compact perinuclear lamellar or vesicular structure characteristic for the Golgi complex (Fig. 18). A weaker and more variable staining of a fine reticular structure, evidently endoplasmatic reticulum, was also detected. When Xcml-GFP-transfected cells were stained with a Golgi-specific dye BODIPY TR ceramide, a significant overlap between green GFP signal and red ceramide signal was observed thereby confirming the preferential localization of the fused protein in the Golgi apparatus. The hydrophobic domain of Xcml is likely to serve as a transmembrane anchor, presumably in a shape of two membrane-spanning a-helices. Deletion of the hydrophobic domain resulted in a marked delocalization of the fused protein, with significant proportion of the signal detected in the

nucleus and cytoplasm (Fig. 18). However, a certain degree of co-localization of the GFP and ceramide signals, although reduced, was still observed. Therefore, it is likely that the hydrophobic domain is essential for the Golgi localization of the Xcml protein; however, it is possibly not the sole targeting signal. As evidenced by injection studies, membrane anchoring seems to be important but not indispensable for Xcml function. The residual activity of mutant protein devoid of membrane anchor might be explained by the part of protein which is still localized to the lumen of secretory pathway organelles, however, more experiments are needed to clarify the issue.

Localization of the Xcml protein to organelles of the secretory pathway suggested the possibility that Xcml might be secreted. To test this, synthetic mRNAs of Xcml with myc epitope tags at the C- or N-terminus were microinjected into Xenopus oocytes followed by Western blotting analysis of the culture medium and vesicular and cytoplasmic fractions of oocytes. Myc tag-containing bands of predicted size were detected only in vesicular fraction (Fig. 18e). When a similar experiment was performed with myc tagged form of secreted protein, immunoreactivity in the culture medium could be easily detected. Hence, Xcml is unlikely to be a secreted protein.

The results of blastomere aggregation assays (Figs. 19 and 20) indicated clearly that Xcml and a human member of the camello family, Hcml1, have a substantial anti-adhesive effect, which confirms earlier data obtained by microscopic observation of Xcml-overexpressing embryos. The blastomere aggregation assay was calcium-based and, therefore, primarily driven by the cadherin adhesion. It has been convincingly demonstrated (Brieher, *et al., J. Cell. Biol.* 126:519-27 (1994); Zhong, *et al., J. Cell. Biol.*, 144:351-59 (1999)) that cadherin C is a major determinant of adhesion in this assay. Therefore, the cadherins, including cadherin C, are the likely targets of Xcml action. Of course, participation of other cell surface or extracellular proteins in Xcml-related anti-adhesive effects remains a definite possibility.

Adhesion is one of the most important mechanisms participating in cancer metastasis, and adhesion proteins (particularly cadherins) have been shown to be important for metastatic processes. Therefore, the camello family proteins, with

their anti-adhesive effects and potential targeting of cadherin-mediated adhesion, are good candidates for the development of anti-metastatic drugs.

It is known that gastrulation movements are maintained by a fine balance of spatially and temporally regulated adhesion. The phenotypes similar to the one produced by overexpression of Xcml can be generated by perturbation of cell adhesion by interference with function of different cadherins or extracellular matrix proteins. Xcml is expressed throughout gastrulation in a critically important region where convergent extension and invagination occur, and its overexpression induces defects similar to those produced by strong reduction of cell adhesion. It is tempting therefore to assume that the normal Xcml function might involve moderate reduction in adhesion of cells located in or moving through the periblastopore region, resulting in change of their migratory properties. This assumption is supported by animal cap elongation experiments which suggest that a controlled reduction of cell adhesion is necessary for gastrulation (Brieher, *et al.*, *J. Cell. Biol.* 126:519-27 (1994)).

Xcml is preferentially localized in Golgi apparatus, which is the major site of synthesis of extracellular matrix proteins as well as terminal processing of cell surface glycoproteins involved in cell adhesion. It is likely that this connection is not coincidental, and that the mechanism of Xcml action may involve participation in the processing of cell surface or extracellular matrix proteins passing through secretory pathway. The strong similarity of Xcml and other members of this family to the two consensus motifs of N-acetyltransferases makes acetylation a natural candidate for this modification. So far, the most prominent acetylation reaction known to occur in Golgi complex is an O-acetylation of sialylic acids in glycoproteins and glycolipids by as yet unidentified enzyme(s). O-acetylation of glycoproteins was shown to change their ashesion to selectins. Whether Xcml may encode sialyc acid O-acetyltransferase remains to be seen, however, the difference between the consensus motifs for N- and O-acetyltransferases does not support this hypothesis.

Camello family can be added to a growing list of proteins such as fringed or Kuzbanian which are localized in Golgi complex and are involved in the regulation embryogenesis. Further, the anti-adhesive effects of camello family

proteins may be implicated in metastasis and tumor aggression, making the proteins an attractive target for anti-metastatic and chemotherapeutic agents.

All publications mentioned hereinabove are hereby incorporated by reference in their entirety.

While the foregoing invention has been described in some detail for purposes of clarity and understanding, it will be appreciated by one skilled in the art from a reading of the disclosure that various changes in form and detail can be made without departing from the true scope of the invention in the appended claims.

What is Claimed is:

1. A purified and isolated nucleic acid encoding a camello protein.

2. The nucleic acid of Claim 1, that is derived from a human, a frog, a mouse, or a rat.

3. The nucleic acid of Claim 2, that is derived from a human.

4. The nucleic acid of Claim 3, which encodes the amino acid sequence for Hcml1 as shown in Figure 2B.

5. The nucleic acid of Claim 4, having the nucleotide sequence for Hcml1 as shown in Figure 2A.

6. The nucleic acid of Claim 3, which encodes the amino acid sequence for Hcml2 as shown in Figure 3B.

7. The nucleic acid of Claim 6, having the nucleotide sequence for Hcml2 as shown in Figure 3A.

8. The nucleic acid of Claim 3, which encodes the amino acid sequence for Hcml3 as shown in Figure 4B.

9. The nucleic acid of Claim 8, having the nucleotide sequence for Hcml3 as shown in Figure 4A.

10. The nucleic acid of Claim 2, which is derived from a frog.

11. The nucleic acid of Claim 10, which encodes the amino acid sequence for Xcml as shown in Figure 1B.

12. The nucleic acid of Claim 11, having the nucleotide sequence for Xcml as shown in Figure 1A.

13. The nucleic acid of Claim 2, which is derived from a mouse.

14. The nucleic acid of Claim 13, which encodes the amino acid sequence for Mcml1 as shown in Figure 5B.

15. The nucleic acid of Claim 14, having the nucleotide sequence for Mcml1 as shown in Figure 5A.

16. The nucleic acid of Claim 13, which encodes the amino acid sequence for Mcml2 as shown in Figure 6B.

17. The nucleic acid of Claim 16, having the nucleotide sequence for Mcml2 as shown in Figure 6A.

18. The nucleic acid of Claim 13, which encodes the amino acid sequence for Mcml3 as shown in Figure 7B.

19. The nucleic acid of Claim 18, having the nucleotide sequence for Mcml3 as shown in Figure 7A.

20. The nucleic acid of Claim 13, which encodes the amino acid sequence for Mcml4 as shown in Figure 8B.

21. The nucleic acid of Claim 20, having the nucleotide sequence for Mcml4 as shown in Figure 8A.

22. The nucleic acid of Claim 13, which encodes the amino acid sequence for Mcml5 as shown in Figure 9B.

23. The nucleic acid of Claim 22, having the nucleotide sequence for Mcml5 as shown in Figure 9A.

24. The nucleic acid of Claim 2, which is derived from a rat.

25. The nucleic acid of Claim 24, which encodes the amino acid sequence for Rcml1 as shown in Figure 10B.

26. The nucleic acid of Claim 25, having the nucleotide sequence for Rcml1 as shown in Figure 10A.

27. The nucleic acid of Claim 24, which encodes the amino acid sequence for Rcml2 as shown in Figure 11B.

28. The nucleic acid of Claim 27, having the nucleotide sequence for Rcml2 as shown in Figure 11A.

29. The nucleic acid of Claim 24, which encodes the amino acid sequence for Rcml3 as shown in Figure 12B.

30. The nucleic acid of Claim 29, having the nucleotide sequence for Rcml3 as shown in Figure 12A.

31. The nucleic acid of Claim 24, which encodes the amino acid sequence for Rcml4 as shown in Figure 13B.

32. The nucleic acid of Claim 31, having the nucleotide sequence for Rcml4 as shown in Figure 13A.

33. A vector comprising nucleic acid encoding a camello protein.

34. A host cell transformed by the vector of Claim 33.

35. A method for producing a recombinant camello protein, comprising growing a host cell transformed with the vector of Claim 33 and isolating the recombinant camello protein from said culture.

36. The method of Claim 35, wherein said host cell is a prokaryotic cell.

37. The method of Claim 35, wherein said host cell is a eukaryotic cell.

38. A purified camello protein or an analogue thereof.

39. The purified camello protein of Claim 38, which is recombinantly produced.

40. A nucleic acid probe which hybridizes to nucleic acid encoding a camello protein.

41. The nucleic acid of Claim 1, having one or more mutations.

42. The nucleic acid of Claim 41, wherein the mutations are selected from the group consisting of a point, insertion rearrangement or deletion mutation.

43. An agent that binds to the protein of Claim 38.

44. The agent of Claim 43, which is an antibody, a peptide, a protein, a nucleic acid, a drug, or antisense nucleic acid.

45. An isolated nucleic acid comprising a nucleotide sequence which is at least 80% homologous with the nucleic acid sequence of Claim 1.

46. An isolated nucleic acid comprising a nucleotide sequence which is at least 85% homologous with the nucleic acid sequence of Claim 1.

47. An isolated nucleic acid comprising a nucleotide sequence which is at least 90% homologous with the nucleic acid sequence of Claim 1.

48. An isolated nucleic acid comprising a nucleotide sequence which is at least 95% homologous with the nucleic acid sequence of Claim 1.

49. An isolated nucleic acid comprising a nucleotide sequence which is at least 98% homologous with the nucleic acid sequence of Claim 1.

50. A non-human, transgenic animal model comprising a nucleic acid encoding camello incorporated into some of the somatic cells of said animal.

51. The animal model of Claim 50, wherein said nucleic acid encodes a functional camello protein.

52. The animal model of Claim 50, wherein said nucleic acid has one or more mutations.

53. An agent that binds to the nucleic acid of Claim 1.

54. An agent that inhibits the expression of the nucleic acid of Claim 1.

55. A method for screening for an agent that binds to the nucleic acid of Claim 1, comprising contacting the nucleic acid with an agent of interest and assessing the ability of the agent to bind to the nucleic acid.

56. A method for screening for an agent that inihibits the expression of the nucleic acid of Claim 1, comprising contacting a cell transformed with a vector comprising the nucleic acid, and assessing the effect of the agent on expression of the nucleic acid.

57. A method for screening for an agent that binds to the protein of Claim 38, comprising contacting the protein with an agent of interest and assessing the ability of the agent to bind to the protein.

Xenopus camello (Xcml)

Nucleotide sequence

```
   1 gcacgagcaa gctgctttct cgttatttct tctgttcccc cggaacagga ctcatataag
  61 atccttctgt agttataggt ggaggccttt gctcagtcgg agtatcatgg ccaacgtctc
 121 cataagaaaa tacaaaaaca gtgactatga aacggtcaac ttcttgtttg ttgaaggaac
 181 aaaagagcat ctcccagcag cctgttggaa cacactgaag aagcctcggt tttatttcat
 241 cattattgtg gcatgtgcca gcatcttcat gtgcaccagt tcctatgttc tgtcccttac
 301 aagccttgtt gccctgttgg ctgttggctg gtatggcttg tacttggaat tccatgggta
 361 tgcaagtcgg tgccagcgtg aggatatgct tgatattgag aattcctaca tgatgagtga
 421 caatacttgt ttctgggtgg cagagataga caggaaggtt gtgggcatag tgggtgccaa
 481 accattaaaa gaagcagatg atgagctgtt tctgttacat ctctctgttg ccagggactg
 541 tcgccagcag cggattggca caaagctgtg ccagacagtc attgattttg ccaggcagcg
 601 tggtttcaaa gctgtgtgtc tggaaacagc aaacatacaa gacgcagcaa taaagttgta
 661 tgaagccgtt ggctttaaga aatcccttgt tgcaatcccc ccattccttc ttaaccaata
 721 cacatctttc acagttattt attacagata tgatatcaaa tcataggaaa tccagtgctt
 781 aataatccat aggacacaat cttctgccac cttccatcag caccggccta cagccacatc
 841 aactggtttc atgagcagaa tcagaaccta agatccaaga tgagtctgaa accctacaga
 901 ctggagaaga ggaaccagtt cagatggtta ttactaaatt cattttggaa agccaccatg
 961 gaaggggaag ctccagaagc ctcctgagat gtttcacttt caatgtcaaa agaaaaataa
1021 acagtagaca aactaatatc aacaagtgtg ggatcgactc tgtccacatg atgtggagta
1081 agaaatttaa ccaatcttaa atcaaagctg ggtatcagtc aatttttctt gattttactc
1141 ttagagtttt ttaaacacag gacatgtcat atgcatttct tctgatattc cttcccatgt
1201 cttgctatta aacagcatat ttgtt
```

FIG. 1A

Predicted amino acid sequence

```
          5         10        15        20        25        30
   1  M A N V S I R K Y K N S D Y E T V N F L F V E G T K E H L P
  31  A A C W N T L K K P R F Y F I I I V A C A S I F M C T S S Y
  61  V L S L T S L V A L L A V G W Y G L Y L E F H G Y A S R C Q
  91  R E D M L D I E N S Y M M S D N T C F W V A E I D R K V V G
 121  I V G A K P L K E A D D E L F L L H L S V A R D C R Q Q R I
 151  G T K L C Q T V I D F A R Q R G F K A V C L E T A N I Q D A
 181  A I K L Y E A V G F K K S L V A I P P F L L N Q Y T S F T V
 211  I Y Y R Y D I K S
```

FIG. 1B

Human camello 1 (Hcml 1)

Nucleotide sequence

```
   1 ccttgggmca gmmttcggca cgagcggcac gagaagcccc agacggtatc tccgagatgc
  61 cagtgagcgg ctgagagctg aagccccctg gacactcaag gctcttgtgg tgacagtctg
 121 acgtaaaggc gtgcagggag gcctagctct gtctcctgga cttagagatt tcagacacag
 181 aagtctgtcc atggctcctt gtcacatccg caaataccag gagagcgacc gccagtgggt
 241 tgtgggcttg ctctcccggg ggatggccga gcatgcccca gccaccttcc ggcaattgct
 301 gaagctgcct cgaaccctca tactcttact tgggggccc ctcgccctac tcctggtctc
 361 tggatcctgg cttctagccc tcgtgttcag catcagcctc ttccctgccc tgtggttcct
 421 tgccaaaaaa ccctggacgg agtatgtgga catgacattg tgcacagaca tgtctgacat
 481 taccaaatcc tacctgagtg agcgtggctc ctgcttctgg gtggctgagt ctgaagagaa
 541 ggtggtgggc atggtaggag ctctgcctgt tgatgatccc accttgaggg agaagcggtt
 601 gcagctgttt catctctctg tggacagtga gcaccgtcgt caggggatag caaaagccct
 661 ggtcaggact gtcctccagt ttgcccggga ccagggctac agtgaagtta tcctggacac
 721 cggcaccatc cagctctctg ctatggccct ctaccagagc atgggcttca agaagacggg
 781 ccagtccttc ttctgtgtgt gggccaggct agtggctctt catacagttc atttcatcta
 841 ccacctccct tcttctaagg tagggagtct gtgatctctt tctgtgtgta ttggtcagaa
 901 tagaatccat tcagctgtag cagcaagcaa tccccaacct ttcactgcaa tgacctttca
 961 atgcaataaa agcttattgt ccattcaaaa aaaaaaaaaa aaaaagatc
```

FIG. 2A

Predicted amino acid sequence

```
          5         10        15        20        25        30
   1  M A P C H I R K Y Q E S D R Q W V V G L L S R G M A E H A P
  31  A T F R Q L L K L P R T L I L L L G G P L A L L L V S G S W
  61  L L A L V F S I S L F P A L W F L A K K P W T E Y V D M T L
  91  C T D M S D I T K S Y L S E R G S C F W V A E S E E K V V G
 121  M V G A L P V D D P T L R E K R L Q L F H L S V D S E H R R
 151  Q G I A K A L V R T V L Q F A R D Q G Y S E V I L D T G T I
 181  Q L S A M A L Y Q S M G F K K T G Q S F F C V W A R L V A L
 211  H T V H F I Y H L P S S K V G S L &
```

FIG. 2B

Human camello 2 (Hcml 2)

Nucleotide sequence

```
   1 ctggactcag tgacttcaga cacagaagtc tgtccatggc tccttatcac atccgcaaat
  61 accaggagag cgaccgcaag tcggtcgtgg gcttgctctc cgggggggatg gccgaacacg
 121 ccccagccac cttccggcga ttactgaagc tgcctcgaac cctcatactc ttacttgggg
 181 gggcccttgc cctactcctg gtctctggct cctggattct ggccctcgtg ttcagcctca
 241 gcctccttcc tgccctgtgg ttccttgcca aaaaaccctg gacgcggtat gtagacatag
 301 cattgcgcac agacatgtct gacatcacca aatcctacct gagtgagtgt ggctcctgct
 361 tctgggtggc tgaatctgaa gagaaggtgg tgggcacagt aggagctctg cccgttgatg
 421 atcccacctt gagggagaag cggttgcagc tgtttcatct ctctgtggac aatgagcacc
 481 gtggtcaggg gatagcaaaa gccctggtca ggactgtcct ccagtttgcc cgggaccagg
 541 gctacagtga agttgtcctg gacaccagca acatccagct ctctgccatg ggcctctacc
 601 agagcttggg cttcaagaag acgggccagt ccttcttcca cgtgtgggcc aggctggtgg
 661 atcttcatac agttcatttc atctatcacc tcccttctgc tcaggcaggg cgtctatgat
 721 ttctttcctt ctgtattggt cagaatagaa tccattcggc tgtagcagca agcaatcccc
 781 aacctctgac tgcaatgacc tttctgtgca ataaaagctt attgtccatt
```

FIG. 3A

Predicted amino acid sequence

```
          5         10        15        20        25        30
   1  M A P Y H I R K Y Q E S D R K S V V G L L S G G M A E H A P
  31  A T F R R L L K L P R T L I L L L G G A L A L L L V S G S W
  61  I L A L V F S L S L L P A L W F L A K K P W T R Y V D I A L
  91  R T D M S D I T K S Y L S E C G S C F W V A E S E E K V V G
 121  T V G A L P V D D P T L R E K R L Q L F H L S V D N E H R G
 151  Q G I A K A L V R T V L Q F A R D Q G Y S E V V L D T S N I
 181  Q L S A M G L Y Q S L G F K K T G Q S F F H V W A R L V D L
 211  H T V H F I Y H L P S A Q A G R L
```

FIG. 3B

Human camello 3 (Hcml 3)

Nucleotide sequence (Partial)

```
   1 gcgctgtgct tcgccgtgag ccgctcgctg ctgctgacgt gcctggtgcc ggccgcgctg
  61 ctgggcctgc gctactacta cagccgcaag gtgatccgcg cctacctgga gtgcgcgctg
 121 cacacggaca tggcggacat cgagcagtac tacatgaagc cgcccggctc ctgcttctgg
 181 gtggccgtgc tggatggcaa cgtggtgggc attgtggctg cacgggccca cgaggaggac
 241 aacacggtgg agctgctgcg gatgtctgtg gactcacgtt tccgaggcaa gggcatcgcc
 301 aaggcgctgg gccggaaggt gctggagttc gccgtggtgc acaactactc cgcggtggtg
 361 ctgggcacga cggccgtcaa ggtggccgcc cacaagctct acgagtcgct gggcttcaga
 421 cacatgggcg cc
```

FIG. 4A

Predicted amino acid sequence

```
          5         10        15        20        25        30
   1  A L C F A V S R S L L L T C L V P A A L L G L R Y Y Y S R K
  31  V I R A Y L E C A L H T D M A D I E Q Y Y M K P P G S C F W
  61  V A V L D G N V V G I V A A R A H E E D N T V E L L R M S V
  91  D S R F R G K G I A K A L G R K V L E F A V V H N Y S A V V
 121  L G T T A V K V A A H K L Y E S L G F R H M G A
```

FIG. 4B

Mouse camello 1 (Mcml 1)

Nucleotide sequence

```
   1 attcggcacg acggctaaaa tggaagtgga gcggactcct agtaccgcta gaagctgctg
  61 gcggaggaca aggagaacta actctaattt gtcccggctt cggaggtgga aaagccccca
 121 ctggtcgggc ctagaagctg agggttcaag gaaggtgtgc aaggcaggta tagctgtctc
 181 tcctggatgc caagatttga gacccagaag tctcccatgg ttccttatca catccgacag
 241 taccaggaca gcgaccataa aagagtcgtg gatgtgttca ccaagggcat ggaggagtac
 301 attccctcta cctttcggca catgcttatg ctgccccgaa ccctcctgct cttacttggg
 361 gtgccccttg ccctggtcct ggtgtctggc tcctggatcc tggctgttat ttgcatcttc
 421 tttctgctcc tacttctgcg gctccttgcc agacagccct ggaaggaata tgtggccaaa
 481 tgtttgcaga cagacatggt tgacatcacc aagtcttacc tgaatgtaca tggcgcctgc
 541 ttctgggtgg ctgagtctgg ggggcaggtg gtgggcatag tggctgctca gccagtcaag
 601 gatcctccac tagggaggaa gcagctgcag ctctttcgcc tgtctgtgtc ctcacagcat
 661 cgaggacagg ggatagcgaa agcgctgacc agaactgtcc tccagtttgc aagggaccag
 721 agttacagtg atgttgtcct tgagaccagc gccttgcagc aaggtgctgt gactctctac
 781 ctgggcatgg gcttcaagaa ggcaggccag tacttcatga gtatattctg gaggttagca
 841 ggtatttgta caattcaatt aaagtactcc ttcccttctg cctaggaggg gtggctgtga
 901 ccttatgctc ctgtgcagca agcacacttc tctgcactct gctacaggaa ccagtgaacc
 961 ctgtcatgtc agtgtgatta acaataaaag ttgttggtgc acaccaaaaa aaaaaaaaaa
1021 aaaaaaa
```

FIG. 5A

Predicted amino acid sequence

```
          5         10        15        20        25        30
   1  M V P Y H I R Q Y Q D S D H K R V V D V F T K G M E E Y I P
  31  S T F R H M L M L P R T L L L L L G V P L A L V L V S G S W
  61  I L A V I C I F F L L L L L R L L A R Q P W K E Y V A K C L
  91  Q T D M V D I T K S Y L N V H G A C F W V A E S G G Q V V G
 121  I V A A Q P V K D P P L G R K Q L Q L F R L S V S S Q H R G
 151  Q G I A K A L T R T V L Q F A R D Q S Y S D V V L E T S A L
 181  Q Q G A V T L Y L G M G F K K A G Q Y F M S I F W R L A G I
 211  C T I Q L K Y S F P S A
```

FIG. 5B

Mouse camello 2 (Mcml 2)

Nucleotide sequence

```
   1 gaggttcacc aggctctggt aggttttact ggatgtcatc ggaggcaaag gccatcctgg
  61 acatttggat ctgtcatatt agactgaatc attccagttg ctggaaagag gatttgttga
 121 aacttggacc tgggaacaca ggagttttca actctgggcc ctgaagagga aacagaagat
 181 ctcagaacag cacatctttc cacagtgtag aacctcagtt cccaaagggc tcagggaagt
 241 tatgcaagaa ggtctggatg tcccttgtga tcactgatac ttgagagcca gaagtctccc
 301 catggctgct tatcacatcc gacagtacca ggagaaggac cacaaaaggg tcctggaatt
 361 gttctccagc ggcatgaagg agcttattcc tgctgccatc cgacagatgc tgacactgcc
 421 tcattctctc ttgctcttac ctggagtgcc tgtgaccata gtattgatgt ctgcctcctg
 481 gctcctggcc acattataca gcttcctctt tctcctttgc ctgtggctta ttttctggat
 541 ttcttgcaga aattatgtgg ctaaaagttt gcaggcagat cttgctgaca tcaccaagtc
 601 ttacctgaat gcacatggct ccttctgggt ggctgagtct ggagaccaag tagttggcat
 661 ggtgggtgct cagccagtca aggaccctcc attagggaag aagcagatgc agctctttcg
 721 cctgtctgtg tcctcacagc atcgaggaca gggaatagca aaggcactgg tcagaactct
 781 cctccagttt gctcgggacc agggttacag tgatgttgtc cttgagactg gcagtgtgca
 841 acatagtgct caggctctct accaggccat gggcttccag aagacaggcc agtactttgt
 901 cagtataagc aagaagttaa tgggtctttc tattcttcaa ttctcttact ctctcccttt
 961 tgcttcagga ccagggtata gtgggaaata tttaaaaaaa ggtcccattc catgctagca
1021 ccaggtactc tctggcccca gtggtctcac tgcctccatg gcttgtccta tgtagcaact
```

FIG. 6A

Predicted amino acid sequence

```
          5         10        15        20        25        30
  1  M A A Y H I R Q Y Q E K D H K R V L E L F S S G M K E L I P
 31  A A I R Q M L T L P H S L L L L P G V P V T I V L M S A S W
 61  L L A T L Y S F L F L L C L W L I F W I S C R N Y V A K S L
 91  Q A D L A D I T K S Y L N A H G S F W V A E S G D Q V V G M
121  V G A Q P V K D P P L G K K Q M Q L F R L S V S S Q H R G Q
151  G I A K A L V R T L L Q F A R D Q G Y S D V V L E T G S V Q
181  H S A Q A L Y Q A M G F Q K T G Q Y F V S I S K K L M G L S
211  I L Q F S Y S L P F A S G P G Y S G K Y L K K G P I P C
```

FIG. 6B

Mouse camello 3 (Mcml 3)

Nucleotide sequence

```
   1 attcggatcc atggcacagc attaaggctg atttggaccc tgagctctga gcaactagtc
  61 taaatgttca gagctgatgg gaaatggctt tgttgaaact tgatcttgga aatcctgcat
 121 ttgcaatgta tatactctag agaaagagat caaaggagct gggcatgaag actggtggcc
 181 tcaagggtta cagggaaacc tacagtcaga agcagctgtg tctttggtct ttgagatctt
 241 agcctccgaa gtctcccatg gctccttatc atatccgaaa ataccaggac agcgaccaca
 301 ggagtgtggt ggatttgttc cgcagaggca tggaggagca catccccgct acctttcgcc
 361 acatgctgct gctgccccga accctcctgc tcttactcgg ggtccctctt actctattcc
 421 tggcctcagg ttcctggctt ctggttcttc tgtccatcct taccctcttt ctttccctgt
 481 ggttccttgc aaaatacaca tgggaaaagc atgtgatgaa ctgtttgcac acagacatgg
 541 ctgacatcac cagaacctac ctgagttctc actcctcctg cttctgggta gctgagtcta
 601 gaggtcagac agtgggcatg gtggctgctc ggccagtgaa ggaccccctc ctgcagaaga
 661 agcaactgca gctacttcac ctctctgtgt cattgcagca ccgaagagaa ggcctaggga
 721 aagctatggt caggactgtc ctccaatttg cacagatgca gggcttcagt gaagttgtcc
 781 tttccaccag catgctgcag tacgcagccc tggctctcta ccagggcatg ggcttccaga
 841 agactggcga gaccttctac acctatttgt ccagactaag gaaatctcca atgataaact
 901 taaagtatag cctcacttct cgggaagggg acctgtga
```

FIG. 7A

Predicted amino acid sequence

```
          5         10        15        20        25        30
   1  M A P Y H I R K Y Q D S D H R S V V D L F R R G M E E H I P
  31  A T F R H M L L L P R T L L L L L G V P L T L F L A S G S W
  61  L L V L L S I L T L F L S L W F L A K Y T W E K H V M N C L
  91  H T D M A D I T R T Y L S S H S S C F W V A E S R G Q T V G
 121  M V A A R P V K D P L L Q K K Q L Q L L H L S V S L Q H R R
 151  E G L G K A M V R T V L Q F A Q M Q G F S E V V L S T S M L
 181  Q Y A A L A L Y Q G M G F Q K T G E T F Y T Y L S R L R K S
 211  P M I N L K Y S L T S R E G D L
```

FIG. 7B

Mouse camello 4 (Mcml 4)

Nucleotide sequence

```
   1 ttcggatcca tgggacactc ggctgtagta gcagctaaga ggacagagag acaagggctg
  61 cgaggcacaa atataaacag atctggtgtc tctcatggat gctgagattt gagacgaagt
 121 ttccccatgg cttcttttcg catccgccag ttccaggaga gggactacaa acaggtcgtg
 181 gatgtgttct ccaggggcat ggaggagcac atacccactg ccttccgcca cttgctgaca
 241 ctgccccgaa ccctcctgct cttagctgtg gtgccccttg ccatagtcct ggtgtctggc
 301 tcctggttcc tggctgttgt atgcattttc tttctgttcc tattcttgtg gttcctcgcc
 361 agcaagccct ggaagaatta tgtgtccaaa tgtttacaca cagacatggc tgacatcacc
 421 aagtcctacc tgagtgtccg tggctcaggt ttctgggtgg ctgagtctgg ggggcaggtg
 481 gtgggtacag tggctgctcg gccagtcaag gatcctccgt tagggaggaa gcagctgcag
 541 ctctttcgcc tgtctgtgtc ctcacagcat cgaggacagg ggatagcgaa agcgctgacc
 601 agaactgtcc tccagtttgc aagggaccag ggttacagtg atgttgtcct tgtgactggc
 661 cttttgcagc aaggtgctgt gactctctac tacagcatgg gcttccagaa gacaggtgaa
 721 tccttcgtgg acatactcac atggcttgtg gatgtttctc taattcattt catataccca
 781 ctcccttctg ctcaaaaata tgagttgtga tctctctcag tgtgtctgtc agcctctggt
 841 ttactatgct gtgggaataa ataacccaga gattgtggtg gacaaatcaa aaaaaaaagg
 901 aaa
```

FIG. 8A

Predicted amino acid sequence

```
          5         10        15        20        25        30
   1  M A S F R I R Q F Q E R D Y K Q V V D V F S R G M E E H I P
  31  T A F R H L L T L P R T L L L L A V V P L A I V L V S G S W
  61  F L A V V C I F F L F L F L W F L A S K P W K N Y V S K C L
  91  H T D M A D I T K S Y L S V R G S G F W V A E S G G Q V V G
 121  T V A A R P V K D P P L G R K Q L Q L F R L S V S S Q H R G
 151  Q G I A K A L T R T V L Q F A R D Q G Y S D V V L V T G L L
 181  Q Q G A V T L Y Y S M G F Q K T G E S F V D I L T W L V D V
 211  S L I H F I Y P L P S A Q K Y E L
```

FIG. 8B

Mouse camello 5 (Mcml 5)

Nucleotide sequence (Partial)

```
  1 caaagtgcta taacctcta tgaggctatg ggattccaaa ggacaggaaa atactcagag
 61 atcagcatta tcaaatggtt aattacattt tctataattc atttcacata ttcttttcct
121 tctactcaga aacatgaact ataatcttat ttcttaccat atagatcagg ttccaattac
181 tgtactgtaa taaataataa aagcatattt ttcatgctca ccggattact acttgacaat
241 gttagggtga caaagttgac ctctacagtg cacagccctt ctccatgaga catttgtttc
301 atctttgaga tcctttccgg gggctacttt gcatctctac tcttattaaa ctgagcat
```

FIG. 9A

Predicted amino acid sequence

```
          5         10        15        20        25        30
  1  Q S A I T L Y E A M G F Q R T G K Y S E I S I I K W L I T F
 31  S I I H F T Y S F P S T Q K H E L
```

FIG. 9B

Rat camello 1 (Rcml 1)

Nucleotide sequence

```
   1 ttcggcacga ggccactgaa tgccactaga agctgatgcc attccagaca ctctaggttg
  61 tgtagtagcg ggactcaggg aaggagtgtg ggcaagtgaa tgctgagatt tgagacccag
 121 aagtttctcc catggtttct tatcacatct gcgagtacca agacagcgac tataaaagtg
 181 ttgtggatgt gtttaccaag ggtgcagaag agtacatccc ctccaccttc cgccacttgc
 241 tgctgctgcc ccgaaccctc ctactcttac ttggggtgtc ccttgccctg gtcctggtgt
 301 ctggctcctg gctgctggct gttgtatgca tctttttttct gctcccattt ttgtggttcc
 361 ttgctggaca gccctggaag aattatgtgt ccaaatgttt acacacagat atggctgaca
 421 tcaccaagtc ttatctgagt gatcgtggct caggtttctg ggtggctgag tctggggagc
 481 aggtagtggg cacagtgggt gctctgccag tcaaggagcc tccatcaggg aggaagcagt
 541 tgcagctctt ccacctggct gtgtcctcac agcatcgagg acaggggata gcgaaagcac
 601 tggtcagaac tgtgctccag tttgcacggg accagggcta cactgatgtt gtccttgaga
 661 ctagcaccat gcagataggt gctgtgaccc tctacctggg catgggtttc cagaagacag
 721 gccaatactt cccgagtatg ctctggaggt tagtgggtat tcgttttgtt caactaaatt
 781 actccttccc ttctgcctag gaagggaggc tgtgaccttg agttcctgtg gagcaagcac
 841 acttccctgc actctgctac aggaaccagt gaaccctgtc atgtcagtgt gattaacaac
 901 aaaagcttgt tgctgc
```

FIG. 10A

Predicted amino acid sequence

```
            5         10        15        20        25        30
  1  M V S Y H I C E Y Q D S D Y K S V V D V F T K G A E E Y I P
 31  S T F R H L L L L P R T L L L L G V S L A L V L V S G S W
 61  L L A V V C I F F L L P F L W F L A G Q P W K N Y V S K C L
 91  H T D M A D I T K S Y L S D R G S G F W V A E S G E Q V V G
121  T V G A L P V K E P P S G R K Q L Q L F H L A V S S Q H R G
151  Q G I A K A L V R T V L Q F A R D Q G Y T D V V L E T S T M
181  Q I G A V T L Y L G M G F Q K T G Q Y F P S M L W R L V G I
211  R F V Q L N Y S F P S A
```

FIG. 10B

Rat camello 2 (Rcml 2)

Nucleotide sequence

```
   1 tcccggcttc ggaagcagaa agcaccctac aggttgggcc tagtagttga gggttcaggg
  61 ataggtatag ctgtctctcc tggatgccaa gatttgagac ccagaagtct cccatggctc
 121 cttatcacat ccgccagtac caagacagcg accacaaaag tgtcgtggat gtgttcacca
 181 agggcatgga agaacacatc ccctccacct tccgccacat gcttatgctg ccccgaaccc
 241 tcctactctt acttggggtg ccccttgccc tggtcctggt gtctggctcc tggctgctgg
 301 ctgttgtatg catcttcttt ctgctcctac tcctgcggtt ccttgctgga cagccctgga
 361 aggagtatgt ggctacatgt ttgcggacag acatggctga catcaccaag tcttacctga
 421 atgcacatgg ctccttctgg gtggctgagt ctggaaacca ggtggtgggc atagtggctg
 481 ctctgccagt caaggatcct ccatcaggga ggaagcagct gcagctcttt cgcctgtctg
 541 tgtcctcaca gcatcgagga caggggatag cgaaagcact ggtcagaact gtcctccagt
 601 ttgcacggga ccagggctac actgatgttg tccttgagac cagtaccttg caacaaggtg
 661 ctatgaccct ctacctgggc atgggcttcc agaagacagg ccaacgcttc ctgactatgt
 721 tctggaggtt agtgggtatt cggacaattc aattaaagta tcccttccct tctgcctagg
 781 aaaggggct gtgaccttga gttcctgtgg agcaagcatg cttctctaaa ctctgctaca
 841 ggaaccagtg aaccctgtca tgtcagtgtg attaacaata aaagcttgtt gctgcacacc
```

FIG. 11A

Predicted amino acid sequence

```
          5         10        15        20        25        30
  1  M A P Y H I R Q Y Q D S D H K S V V D V F T K G M E E H I P
 31  S T F R H M L M L P R T L L L L L G V P L A L V L V S G S W
 61  L L A V V C I F F L L L L L R F L A G Q P W K E Y V A T C L
 91  R T D M A D I T K S Y L N A H G S F W V A E S G N Q V V G I
121  V A A L P V K D P P S G R K Q L Q L F R L S V S S Q H R G Q
151  G I A K A L V R T V L Q F A R D Q G Y T D V V L E T S T L Q
181  Q G A M T L Y L G M G F Q K T G Q R F L T M F W R L V G I R
211  T I Q L K Y P F P S A
```

FIG. 11B

Rat camello 3 (Rcml 3)

Nucleotide sequence (Partial)

```
   1 tgtcaggcca agaatttcggc acgaggagga cagcgaccac aggagtgtag tgaatttgtt
  61 ctgcagaggc acggaggagc acatctccgc cagcttccgc tacatgctgc tgctgcccgg
 121 aaccctcctg atcttactcg gggtccctct tactctattc ttggcctcag gctcctggct
 181 tctggttctt ctgtccaccc taaccctcct tgtttccctg tggctccttg caaaataccc
 241 ttgggagaag tatacggcaa tgtgtttgca ctcagacatg gctgatatcc ccagaaccta
 301 cttgagttct cattactcct gcttctgggt ggctgagtct agaggtcaga tggtgggcat
 361 aatcgctgtt ttaccagtga aggatcccct cctgcagagg aagcaactgc agctacgtca
 421 cctctctgtg tccctggagc accggagaga ggggattgga agagctatgg tcaggactgc
 481 cctccagttt gcagagatgc agggcttcag tgaagttgtc ctggtcacca gcatgttgca
 541 gtatgctgcc ctagctctgt accagagcat gggcttccag aagactggtg agttcttcta
 601 taccttttgtc tctcgactaa ggaattctcc aatgatatgc ttaaaatatt gcctcacttc
 661 tgctctgaat gacctgaaaa cctgaaagac ctgctctgag agacctgtga gctctctcct
 721 gtggccatca gtcaggatct aattgcttct gtaatagtaa caagcaaacc cagctatttc
 781 agcaaaccac tgaccctcac tctcaagcac atcggaataa atgtttgtgg atggggttgg
 841 ggcaatggct actctttgtt atccatgctt ttctgaggta tcctttagct aatactacaa
 901 tcatatataa aaagtaacgc agataataaa atttaactta gcttgtg
```

FIG. 12A

Predicted amino acid sequence

```
          5         10        15        20        25        30
   1  M V R P R I R H E E D S D H R S V V N L F C R G T E E H I S
  31  A S F R Y M L L L P G T L L I L L G V P L T L F L A S G S W
  61  L L V L L S T L T L L V S L W L L A K Y P W E K Y T A M C L
  91  H S D M A D I P R T Y L S S H Y S C F W V A E S R G Q M V G
 121  I I A V L P V K D P L L Q R K Q L Q L R H L S V S L E H R R
 151  E G I G R A M V R T A L Q F A E M Q G F S E V V L V T S M L
 181  Q Y A A L A L Y Q S M G F Q K T G E F F Y T F V S R L R N S
 211  P M I C L K Y C L T S A L N D L K T
```

FIG. 12B

Rat camello 4 (Rcml) 4)

Nucleotide sequence

```
   1 agacgaaggt ttcccatggc ttcttttcac atccgccagt tccaggagag ggactatgaa
  61 caggtcgtgg atatgttctc caggggaatg aaggaacaca tccccactgc cttccgccac
 121 ttgctgctgc tgccccgaac cctcctactc ttacttgggg tgccccttgc cctggtcctg
 181 gtgtctggct cctggctgct ggctgttgta tgcatcttct ttctgctccc atttttgtgg
 241 ttccttgctg gacagccctg gaagaattat gtgtccaaat gcttacacac agacatggct
 301 gacatcacca agtcttatct gagtgatcgt ggctcaggtt tctgggtggc tgagtctggg
 361 ggccagatag tgggcacagt gggtgctctg ccagtcaagg atcctccatc agggaggaag
 421 cagttgcagc tcttccgcct gtctgtgtcc tcacagcatc gaggacaggg gatagcgaaa
 481 gcactggtca gaactgtgct ccagtttgca cgggaccagg gctacacgga tgttgtcctt
 541 gtgactggcc ttttgcagca aggtgctgtg accctctact acagcatggg cttccagaag
 601 acaggcgaat ccttcatgga catactcaca tggcttgtgg atgtttctct aattcatttc
 661 atatacccgc tcccttcctc ctgagaacct gagtttcgat ccctctgtgt gtctgtcagc
 721 ctctggttca ctgtgctgtg ggaacaaata atcctgatat tgtagtggac aaatcaccc
```

FIG. 13A

Predicted amino acid sequence

```
          5         10        15        20        25        30
   1  M A S F H I R Q F Q E R D Y E Q V V D M F S R G M K E H I P
  31  T A F R H L L L L P R T L L L L L G V P L A L V L V S G S W
  61  L L A V V C I F F L L P F L W F L A G Q P W K N Y V S K C L
  91  H T D M A D I T K S Y L S D R G S G F W V A E S G G Q I V G
 121  T V G A L P V K D P P S G R K Q L Q L F R L S V S S Q H R G
 151  Q G I A K A L V R T V L Q F A R D Q G Y T D V V L V T G L L
 181  Q Q G A V T L Y Y S M G F Q K T G E S F M D I L T W L V D V
 211  S L I H F I Y P L P S S
```

FIG. 13B

Alignment of amino acid sequences of camello protein family members

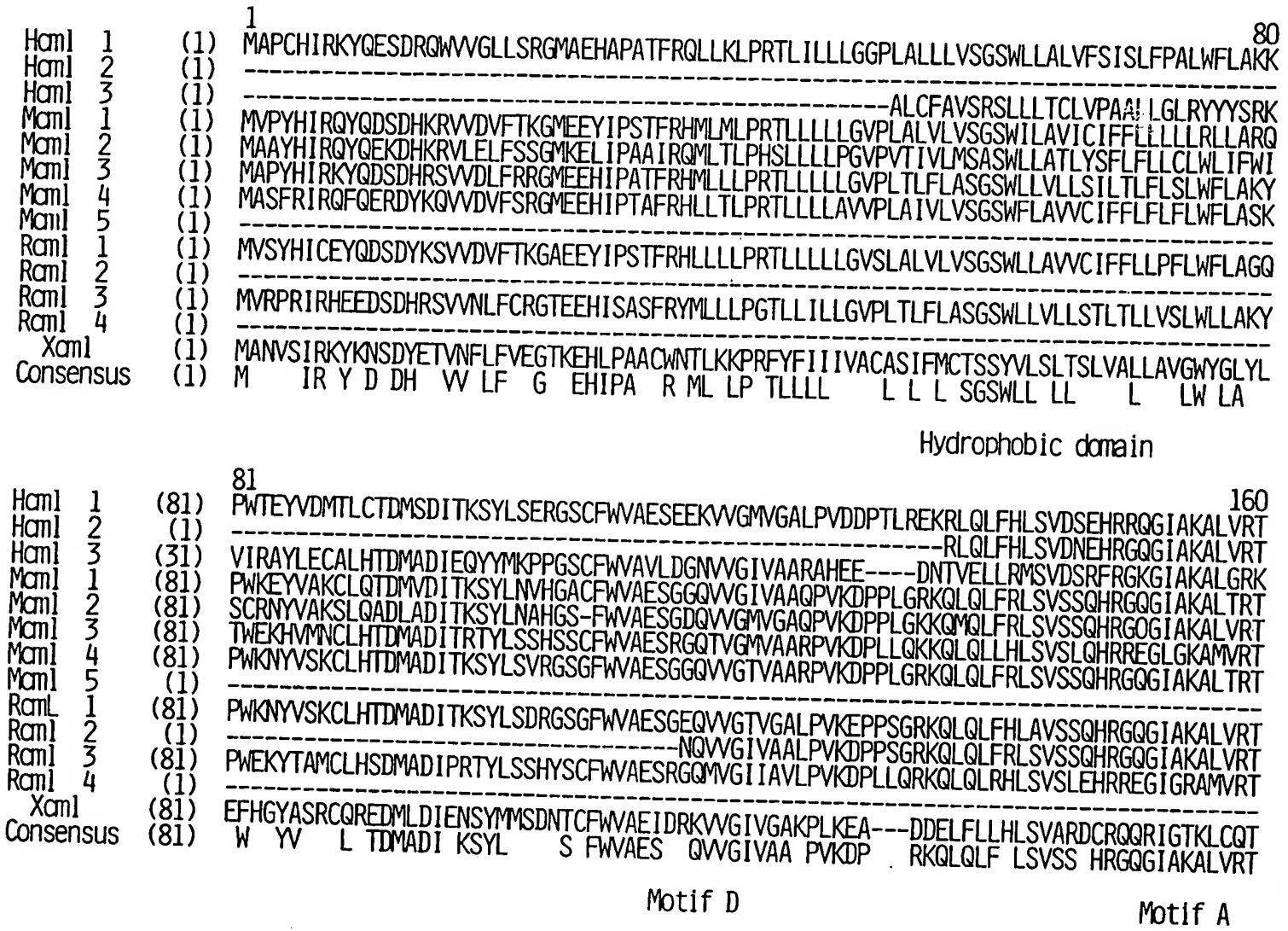

FIG. 14-1

```
                      161                                                                            240
Haml 1      (161)  VLQFARDQGYSEVILDTGTIQLSAMALYQSMGFKKTGQS-FFCVWARLVALHTVHFIYHLPSSKVGSL------------
Haml 2       (26)  VLQFARDQGYSEWLDTSNIQLSAMGLYQSLGFKKTGQS-FFHVWARLVDLHTVHFIYHLPSAQAGRL------------
Haml 3      (107)  VLEFAVVHNYSAVVLGTTAVKVAAHKLYESLGFRHMGA-----------------------------------------
Maml 1      (161)  VLQFARDQSYSDVVLETSALQQGAVTLYLGMGFKKAGQY-FMSIFWRLAGICTIQLKYSFPSA----------------
Maml 2      (160)  LLQFARDQGYSDVVLETGSVQHSAQALYQAMGFQKTGQY-FVSISKKLMGLSILQFSYSLPFASGPGYSGKYLKKGPIPC
Maml 3      (161)  VLQFAQMQGFSEVVLSTSMLQYAALALYQGMGFQKTGET-FYTYLSRLRKSPMINLKYSLTSREGDL------------
Maml 4      (161)  VLQFARDQGYSDVVLVTGLLQQGAVTLYYSMGFQKTGES-FVDILTWLVDVSLIHFIYPLPSAQKYEL-----------
Maml 5        (1)  --------------------QSAITLYEAMGFQRTGKYSEISIIKWLITFSIIHFTYSFPSTQKHEL-----------
Raml 1      (161)  VLQFARDQGYTDVVLETSTMQIGAVTLYLGMGFQKTGQY-FPSMLWRLVGIRFVQLNYSFPSA----------------
Raml 2       (46)  VLQFARDQGYTDVVLETSTLQQGAMTLYLGMGFQKTGQR-FLTMFWRLVGIRTIQLKYPFPSA----------------
Raml 3      (161)  ALQFAEMQGFSEVVLVTSMLQYAALALYQSMGFQKTGEF-FYTFVSRLRNSPMICLKYCLTSA-LNDLKT---------
Raml 4        (1)  --------GYTDVVLVTGLLQQGAVTLYYSMGFQKPGES-FMDILTWLVDVSLIHFIYPLPSS----------------
Xaml        (158)  VIDFARQRGFKAVCLETANIQDAAIKLYEAVGFKKSLVAIPPFLLNQYTSFTVIYYRYDIKS----------------
Consensus          VLQFARDQGYSDVVLETS LQ GAV LY SMGFQKTG   F SIL RLV I II F Y LPSA
```

Motif B

FIG. 14-2

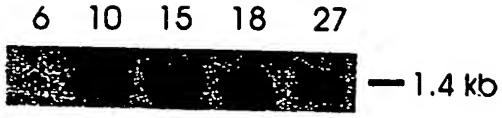

FIG. 15A

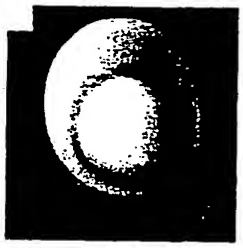

FIG. 15B

FIG. 15C

FIG. 15D

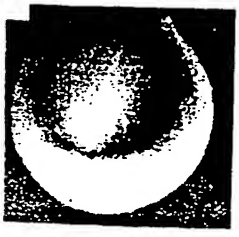

FIG. 15E

FIG. 15F

FIG. 15G

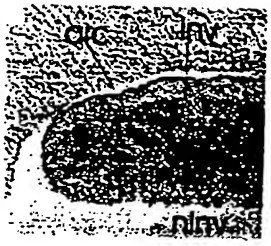

FIG. 15H

FIG. 16A

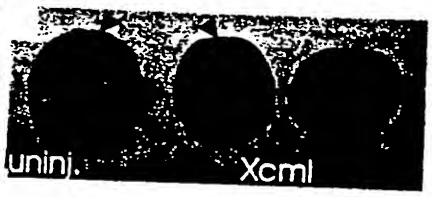

FIG. 16B

FIG. 16C

FIG. 16D

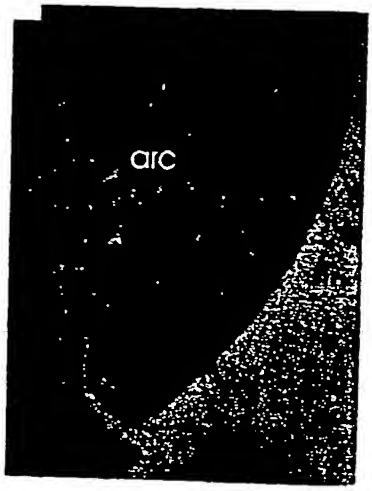

FIG. 16E

FIG. 16F

FIG. 16G

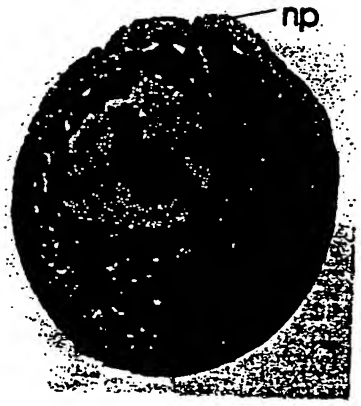

FIG. 16H

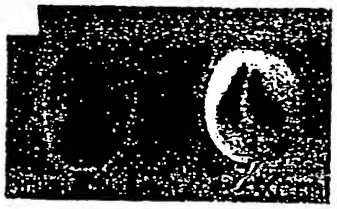

FIG. 17A

FIG. 17B

FIG. 17C

FIG. 17F

FIG. 17D

FIG. 17E

FIG. 17G

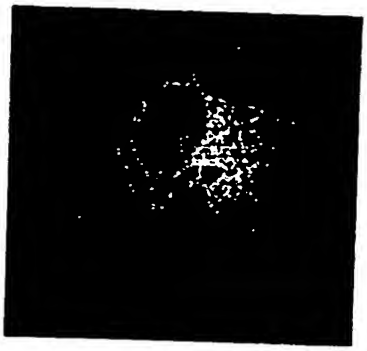

FIG. 18A

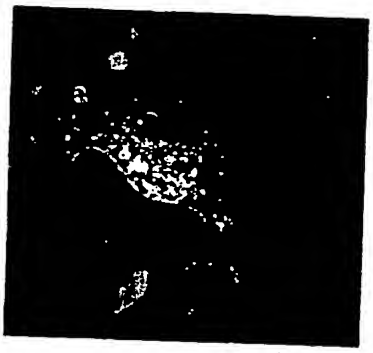

FIG. 18B

FIG. 18C

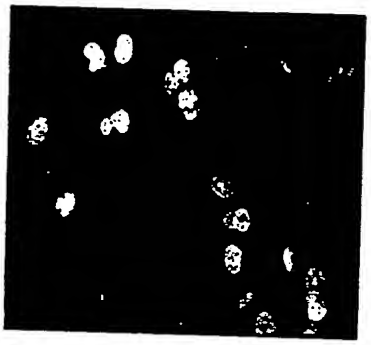

FIG. 18D

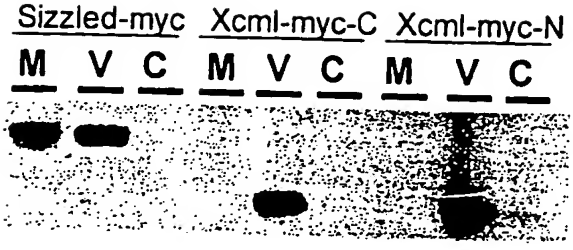

FIG. 18E

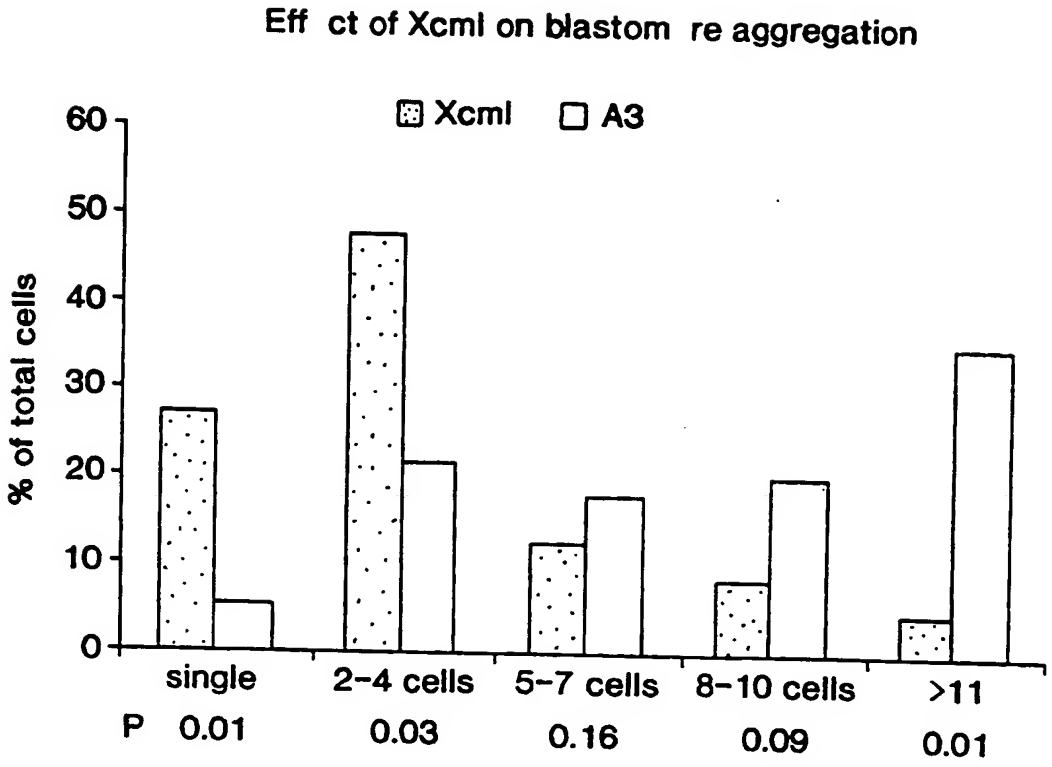
Eff ct of Xcml on blastom re aggregation
Xcml
A3
% of total cells
60
50
40
30
20
10
0
single
2-4 cells
5-7 cells
8-10 cells
>11
P 0.01
0.03
0.16
0.09
0.01

FIG. 19

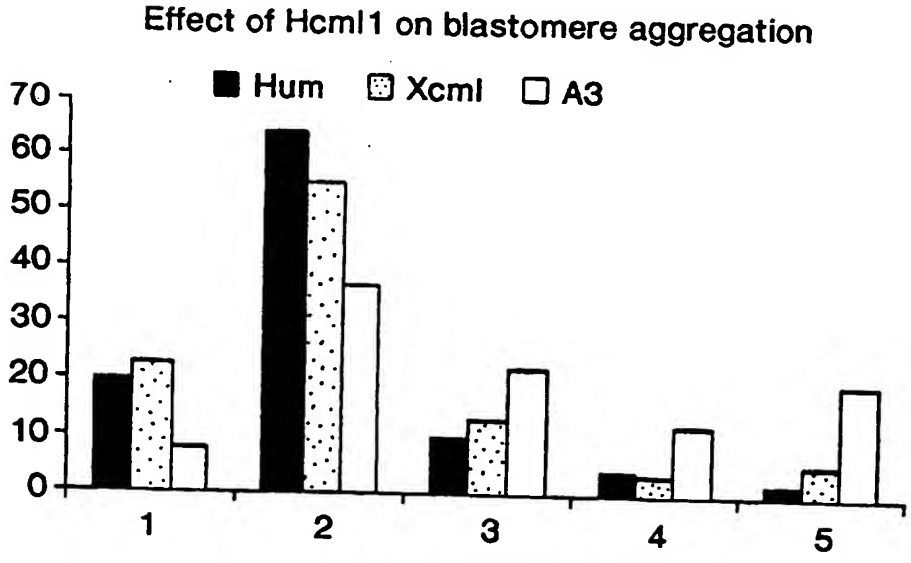
Effect of Hcml1 on blastomere aggregation
Hum
Xcml
A3
70
60
50
40
30
20
10
0
1
2
3
4
5

FIG. 20

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/16412

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) :C07H 21/02, 21/04; A01K 67/00
US CL : 536/23.1; 800/8
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.1; 800/8

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

STN, MEDLINE, EAST, NCBI database

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	POPSUEVA et al. Camello, a novel gene involved in regulation of xenopus gastrulation. Developmental Biology. 01 June 1999, Vol. 210, No. 1, page 235, abstract 326, see entire abstract.	1-3, 10, 13, 24, 33-49
X	IVANOVA et al. Identification of mRNA, localized at various segments of the Xenopus laevis embryo at early stages of the gastrula. Dokl Akad Nauk. March 1998, Vol. 359, No. 1, pages 116-119, see entire document.	1-3, 10, 13, 24, 33-49

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
23 AUGUST 2000	30 AUG 2000
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer JOSEPH T. WOITACH Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet) (July 1998)★